OM nucleic - nucleic search, using sw model

Run on: May 27, 2003, 11:58:11 ; Search time 7222 Seconds
                                        (without alignments)
                                        11097.913 Million cell updates/sec

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

Title: US-09-808-743A-1
Perfect score: 2754
Sequence: 1 atgatcgcctcgcatatgat...........gggaggctggcagagatag 2754
Scoring table: OLIGO_NUC
                 Gapop 60.0 , Gapext 60.0
Searched: 2054640 seqs, 145514027878 residues
Word size :   50
Total number of hits satisfying chosen parameters: 18
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   GenEmbl:*

        1: gb_ba:*
        2: gb_htg:*
        3: gb_in:*
        4: gb_om:*
        5: gb_ov:*
        6: gb_pat:*
        7: gb_ph:*
        8: gb_pl:*
        9: gb_pr:*
       10: gb_ro:*
       11: gb_sts:*
       12: gb_sy:*
       13: gb_un:*
       14: gb_vi:*
       15: em_ba:*
       16: em_fun:*
       17: em_hum:*
       18: em_in:*
       19: em_mu:*
       20: em_om:*
       21: em_or:*
       22: em_ov:*
       23: em_pat:*
       24: em_ph:*
       25: em_pl:*
       26: em_ro:*
       27: em_sts:*
       28: em_un:*
       29: em_vi:*
       30: em_htg_hum:*
       31: em_htg_inv:*
       32: em_htg_other:*
       33: em_htg_mus:*
       34: em_htg_pln:*
       35: em_htg_rod:*
       36: em_htg_man:*
       37: em_htg_vrt:*
       38: em_sy:*
       39: em_htgo_hum:*
       40: em_htgo_mus:*
       41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2754	100.0	2754	10	AF027179	AF027179 Rattus no
2	2550	92.6	3635	6	AR068839	AR068839 Sequence
3	2550	92.6	3635	6	AR070139	AR070139 Sequence
4	2550	92.6	3635	10	RATHKII	RATHKII Rat hexokin
5	1450	52.7	2236	10	RATHKIINAH	M68971 Rat hexokin
6	305	11.1	2317777	2	AC098769	M68972 Rat hexokin
7	187	6.8	2317777	2	AC098769	AC098769 Rattus no
8	95	3.4	5474	10	MMU238540	AC098769 Rattus no
9	83	3.0	28646	10	MMHEXII3	AJ238540 Mus muscu
10	83	3.0	67493	2	AC116811	Y11668 Mus musculi
11	83	3.0	214839	2	AC007305	AC116811 Mus muscu
12	72	2.6	404	10	MMHEXII2	AC007305 Mus muscu
13	65	2.4	1731	10	RATHK2	Y11667 Mus musculu
14	65	2.4	5067	10	AY082375	D26393 Rattus norv
15	65	2.4	5150	10	RNU19605	AY082375 Rattus no
16	65	2.4	6151	10	MMHEXIII1	U19605 Rattus norv
17	50	1.8	200	10	MMDEFI	Y11666 Mus musculu
18	50	1.8	67493	2	AC116811	X96699 M.musculus
						AC116811 Mus muscu

ALIGNMENTS

RESULT 1
AF027179
LOCUS   AF027179   2754 bp   mRNA   linear   ROD 16-DEC-1997
DEFINITION   Rattus norvegicus mutant type II hexokinase mRNA, complete cds.
ACCESSION   AF027179
VERSION   AF027179.1 GI:2689657
KEYWORDS   .
SOURCE   Rattus norvegicus.
  ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE   1 (bases 1 to 2754)
  AUTHORS   Mathupala,S.P.
  TITLE   Type II Hexokinase of Hepatoma AS-30D
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 2754)
  AUTHORS   Mathupala,S.P.
  TITLE   Direct Submission
  JOURNAL   Submitted (27-SEP-1997) Biological Chemistry, Johns Hopkins
             University School of Medicine, 725 N. Wolfe Street, Baltimore, MD
             21205, USA
FEATURES   Location/Qualifiers
     source   1..2754
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /cell_line="AS-30D"
                     /cell_type="ascites hepatoma"
     CDS   1..2754
                     /note="contains 4 amino acid mutations when compared to
                     skeletal muscle type II hexokinase"
                     /codon_start=1
                     /product="mutant type II hexokinase"
                     /protein_id="AAB91396.1"
                     /db_xref="GI:2689658"
                     /translation="MIASHMIACLFTELNQNQVQKVDQFLYHMRLSDETLLEISRRFR
                     KEMEKGLGATTHPTAAVKMLPTFVRSTPDGTEHGEFLALDLGGTNFRVLRVRVTDNGL
                     QRVEMENQIYALEDIMRGSGTQLFDHIAECLANFMDKLQIKEKKLPLGFTSFPCHQ
                     TKLDESFLVSWTKGFKSSGVEGRDVVDLIRKVIQRGDFDIDIVAVVNDTVGTMMTCG
                     YDDQNCEIGLIVGTGSNACYVEEMRHIDMVEGDEGRMCINMEWGAFGDGTLNDIRTE
                     FDREIDMGSLNPGKQLFEKMISGMYMGELVRLILVKMAKAELFQGKLSPELLTTGSF"

```
          ETKDVSDIEEDKDGIEKAYQILMRLGLNPLQEDEDGVATHRICQIVSTRSASLCAATLAA
          VLWRIKENKGEERLRSTIGVDGSVYKKHPHPFAKIRLHKAVRRLVPDCDVRFLRSEDGSG
          KGAAWTAVAYRLADQHRAROKRTLESLKLSHEQLLEVKRRMKVEMBEQGLSKETHAVAP
          VKMLPTYVCATPDGTEKGDFLALDLGGTNFRVLLVRVRNGKRRGVEMHNKIYSIPQEV
          MHGTGEELFDHIVQCIATDFLEYMGMKGVSLPLGFTFSFPCQQNSLDQSILLKWTKGFK
          ASGCEGEDVVTLLKEAIHRREEFDLDVVAVVNDTVGTMMTCGYEDPHCEVGLIVGTGS
          FEKMISGMYLGEIVRNLLIDLKTRGLLFRGRISERLKTRGISEFKLSQIESDCLALL
          NACYMEEMRNVELVDGEEGRMCVNMEWGAFGDNGCLDDLRTVEFDVAVDELSLNPGKOR
          QVRAILRHLIGLESTCDDSIIVKEVCTVVARRAAQLCGAGMAAVVDKIRENRGLDHPKV
          TVGVDGTLYKLHPHFAKVMHETVRDLAPKCDVSFLESEDGSGKGAALITAVACRIREA
          GQR"

variation      344
               /note="nucleotide change resulting in an amino acid change
               of Leu instead of Pro"
               /replace="c"

variation      575
               /note="nucleotide change resulting in an amino acid change
               of Val instead of Ala"
               /replace="c"

variation      2345
               /note="nucleotide change resulting in an amino acid change
               of Ser instead of Phe"
               /replace="t"

variation      2561
               /note="nucleotide change resulting in an amino acid change
               of Pro instead of Leu"
               /replace="a"

BASE COUNT    642 a   659 c   869 g   584 t
ORIGIN

Query Match          100.0%;   Score 2754;   DB 10;   Length 2754;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 2754;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
```

```
Qy     1 ATGATCGCCTCGCAATATGATGATCCGTGGTTGAATGACACAGTTGGGACCAATGAT    60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGATCGCCTCGCAATATGATGATCCGTGGTTGAATGACACAGTTGGGACCAATGAT    60

Qy    61 AAGGTTCACCAATTTCTCTACCACATGCGTCTCAGATGATGAGACCCTTCTGGAGATTTCT   120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AAGGTTCACCAATTTCTCTACCACATGCGTCTCAGATGATGAGACCCTTCTGGAGATTTCT   120

Qy   121 AGGCGGTTCCCGGAAGGAGAATGGAGAAAGGGCTAGGAGCTACCACGGCCGACCCTTCT   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGGCGGTTCCCGGAAGGAGAATGGAGAAAGGGCTAGGAGCTACCACGGCCGACCCTTCT   180

Qy   181 GTGAAAAATGTTGCCTACCTTTGTGAGGTCACACTGCCGGACAGAGAACATGGGAGTTC   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGAAAAATGTTGCCTACCTTTGTGAGGTCACACTGCCGGACAGAGAACATGGGAGTTC   240

Qy   241 CTGGCTCTGGATCTTGGAGGAACCACAACTTCCGTGTGCTCCGAGTAAGGGTGACGGACAAT   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTGGCTCTGGATCTTGGAGGAACCACAACTTCCGTGTGCTCCGAGTAAGGGTGACGGACAAT   300

Qy   301 GGCCTCCAGAGAGTGGAGATGGAAGAACCAGAACTACGCCCATCCTCTTGAGGACACATCG   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGCCTCCAGAGAGTGGAGATGGAAGAACCAGAACTACGCCCATCCTCTTGAGGACACATCG   360

Qy   361 GGCAGTGGAACCCAGCTGTTTGACCAGCCGCCGAATGCCTGGCCATTCATTGGACAAG   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCAGTGGAACCCAGCTGTTTGACCAGCCGCCGAATGCCTGGCCATTCATTGGACAAG   420

Qy   421 CTACAAAATCAAAGAGAAGAGAAGAAAGCTCCCCTCTGGGTTTCACTTTCTCGTTCCCCTG   480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTACAAAATCAAAGAGAAGAGAAGAAAGCTCCCCTCTGGGTTTCACTTTCTCGTTCCCCTG   480

Qy   481 ACAAAAACTGGATGACAGAGTTTTTTGGTCTCGTGGACTAAGGGTTCAAGTCCACTGGCGTG   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACAAAAACTGGATGACAGAGTTTTTTGGTCTCGTGGACTAAGGGTTCAAGTCCACTGGCGTG   540

Qy   541 GAAGGCAGAGAGATGTGTGGTGGACCTGATCCGGAGAAGGTTATCCAGCGGAGAGGGACTTTGAC   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GAAGGCAGAGAGATGTGTGGTGGACCTGATCCGGAGAAGGTTATCCAGCGGAGAGGGACTTTGAC   600
```

```
Qy   601 ATTGACATTGTGGCCCGTGGTTGAATGACACAGTTGGGACCAATGATGACTTGTGGCTATGAT   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATTGACATTGTGGCCCGTGGTTGAATGACACAGTTGGGACCAATGATGACTTGTGGCTATGAT   660

Qy   661 GATCAGAAACTCGCAGAGATTGGTCTCATTGTGGGCACTGGCAGCAAGGCCTTGCTCACGGGAG   720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GATCAGAAACTCGCAGAGATTGGTCTCATTGTGGGCACTGGCAGCAAGGCCTTGCTCACGGGAG   720

Qy   721 GAAAATGCGTCAATATTGACATGGTGGGAGGAGGAGATGAGGGGCCATGTGCATCAACACATGGGAG   780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAAAATGCGTCAATATTGACATGGTGGGAGGAGGAGATGAGGGGCCATGTGCATCAACACATGGGAG   780

Qy   781 TGGGGAGCCCTTTGGGGGGACGAGGGTACACTCAATGACATGACTTGTTGACCGAGAG   840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TGGGGAGCCCTTTGGGGGGACGAGGGTACACTCAATGACATGACTTGTTGACCGAGAG   840

Qy   841 ATGGACATTGGGGCTGCTGAACCCTGGGAAGCAGCAGCAGGTTGGAGAAGATGAATTAGCGGGGATG   900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 ATCGACAACTCGGAGAGGTGGGTCGCGTGAACCCTGGGAAGCAGCAGCAGGTTTGAGAAGATGAATTAGCGGGGATG   900

Qy   901 TACATGGGGGAGGTCAGGGTGGTGGCTCATCCTGGTGAAGATGGGCCAAGGCCAGAGACAGAGACAT   960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TACATGGGGGAGGTCAGGGTGGTGGCTCATCCTGGTGAAGATGGGCCAAGGCCAGAGACAGAGACAT   960

Qy   961 CAACGGGAAACTCACCCACAGAAACTCCTTACCACTGGCCTCCTTCGAGACCAAAGATGTCTCG   1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CAACGGAAACTCACCCACAGAAACTCCTTACCACTGGCCTCCTTCGAGACCAAAGATGTCTCG   1020

Qy  1021 GATATTGAAGAGGATAAGGATGAAATCGAGAAGGCCTACAATGACATCCAAATCCTGATTGGGCCGTGGT   1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 GATATTGAAGAGGATAAGGATGAAATCGAGAAGGCCTACAATGACATCCAAATCCTGATTGGGCCGTGGT   1080

Qy  1081 CTGAATCCATTGCAGGAGGATTGTGGTGGCCACGCACCGGAGAAGTCCCAGATTGTGTCCACG   1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 CTGAATCCATTGCAGGAGGATTGTGGTGGCCACGCACCGGAGAAGTCCCAGATTGTGTCCACG   1140

Qy  1141 CGCTCGGCCCAGTCGTGTCGCGGCAGCCACCATCGCCGGGGTCGTGGCGGAATCAAAGAGAAC   1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CGCTCGGCCCAGTCGTGTCGCGGCAGCCACCATCGCCGGGGTCGTGGCGGAATCAAAGAGAAC   1200

Qy  1201 AAGGGCGGAGGAGGAGGCGGACTTCCGTCACCACCATCGGGTTCGGATGGGCACCGGCAGTGAGT   1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 AAGGGCGGAGGAGGAGGCGGACTTCCGTCACCACCATCGGGTTCGGATGGGCACCGGCAGTGAGT   1260

Qy  1261 CCCATTTTGCCAAGCGTCTCCATAAGGCAGTGAGGAGGCTGGTGCCCGACTGTGATGTC   1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 CCCATTTTGCCAAGCGTCTCCATAAGGCAGTGAGGAGGCTGGTGCCCGACTGTGATGTC   1320

Qy  1321 CGCTTCCTCCGCTCTGAGGATGGCAGCGGCAAGGGCGCGGCT   1380
         ||||||||||||||||||||||||||||||||||||||||||
Db  1321 CGCTTCCTCCGCTCTGAGGATGGCAGCGGCAAGGGCGCGGCT   1380

Qy  1381 TACCGTCTGGCTGACCAACACACCGGCCCAGAGAGAGTCTCGAAGCTGAGC   1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 TACCGTCTGGCTGACCAACACACCGGCCCAGAGAGAGTCTCGAAGCTGAGC   1440

Qy  1441 CACGAGCAGCGCTTCTGAGCGTTAAGAAGAAGAATGAAGGTGGAAATGAGAC   1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 CACGAGCAGCGCTTCTGAGCGTTAAGAAGAAGAATGAAGGTGGAAATGAGAC   1500

Qy  1501 AAGGAGACGGACCATGCGGTCGCCCCTGAAGCTGTCAAGAAGATGAAGGTGTGTGTGCATCCA   1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 AAGGAGACGGACCATGCGGTCGCCCCTGAAGCTGTCAAGAAGATGAAGGTGTGTGTGCATCCA   1560

Qy  1561 GATGGCACAGAAAAGGAGGACTTGCCCTGAGCTTGGGATCTTGGACCTGGGTATGGAATTGCTAC   1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 GATGGCACAGAAAAGGAGGACTTGCCCTGAGCTTGGGATCTTGGACCTGGGTATGGAATTGCTAC   1620

Qy  1621 CTGCTGGTGCGTGTGCGTAATGGCAAGCGGCGTGGAGTGCATGAACATAACAAAGATCTAC   1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CTGCTGGTGCGTGTGCGTAATGGCAAGCGGCGTGGAGTGCATGAACATAACAAAGATCTAC   1680
```

```
RESULT 2
LOCUS       AR068839                3635 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 15 from patent US 5854067.
ACCESSION   AR068839
VERSION     AR068839.1  GI:6001046
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3635)
  AUTHORS   Newgard,C.B., Han,H.-P. and Normington,K.D.
  TITLE     Hexokinase inhibitors
  JOURNAL   Patent: US 5854067-A 15 29-DEC-1998;
FEATURES             Location/Qualifiers
     source          1..3635
                     /organism="unknown"
BASE COUNT      816 a    875 c   1098 g    846 t
ORIGIN

Query Match          92.6%;  Score 2550;  DB 6;  Length 3635;
Best Local Similarity   99.9%;
Matches 2750;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy      1  ATGATCGCCTCGCATATGATCGCCGCTGCTTATTCACGGAGCTCAACAAGTGCAG    60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198  ATGATCGCCTCGCATATGATCGCCGCTGCTTATTCACGGAGCTCAACAAGTGCAG   257

Qy     61  AAGGTTGACCAATTTCTCTCACCACATGCGTCTCTCAGATAGAGACCCTTCTGGAGATTTCT   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258  AAGGTTGACCAATTTCTCTCACCACATGCGTCTCTCAGATAGAGACCCTTCTGGAGATTTCT   317

Qy    121  AGGCCGGTTCCGGAAGGAGATGGAGAAAGGGCTAGGAGACTACCAGCCACCCTACAGCACGT   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318  AGGCCGGTTCCGGAAGGAGATGGAGAAAGGGCTAGGAGACTACCAGCCACCCTACAGCACGT   377

Qy    181  GTGAAAATGTTGCCCTACCTTTGTGAGGTCAACTCCGGATGGGACAGAACATGGGGAGTTC   240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    378  GTGAAAATGTTGCCCTACCTTTGTGAGGTCAACTCCGGATGGGACAGAACATGGGGAGTTC   437

Qy    241  CTGGCCTCGGATCTTGGAGGAGGAACCCAACTCACCTCCGGTCGTCCGAGTGGTGACGGACAAT   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    438  CTGGCTCTCGGATCTTGGAGGAGGAACCCAACTCACCTCCGGTCGTCCGAGTGGTGACGGACAAT   497

Qy    301  GGCCTCCAGAGAGTGGAGATGGAGAACCAGATCTACGGCCATCCTGAGGACACATCATGCGG   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    498  GGCCTCCAGAGAGTGGAGATGGAGAACCAGATCTACGGCCATCCTGAGGACACATCATGCGG   557

Qy    361  GGCAGTGGAACCAGCCTGTTTTGACCACATCGCGAAATGCCTGCCAACTTCATGGACAAG   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    558  GGCAGTGGAACCAGCGTGTTTGACCACATCGCGAAATGCCTGCCAACTTCATGGACAAG   617

Qy    421  CTACAAATCAAAGAGAGAGAAGCTCCCTCTGGGTTTCACCTTCTCGTTCCCTGCCGCGCC   480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    618  CTACAAATCAAAGAGAGAGAAGCTCCCTCTGGGTTTCACCTTCTCGTTCCCTGCCGCGCC   677

Qy    481  ACAAAACTGGATGAGAGTTTTTTGGTCTCGTGGGACTAAGGTGGCTGCCGTG   540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    678  ACAAAACTGGATGAGAGTTTTTTGGTCTCGTGGGACTAAGGGTTCAAGTCCAGTGGCCGTG   737

Qy    541  GAAGGCAGAGATGGTGGTGGACCTGATCCGGAAGGTTATCCAGCCGGAGAGGGGGACTTCGAC   600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    738  GAAGGCAGAGATGGTGGTGGACCTGATCCGGAAGGTTATCCAGCCGGAGAGGGGGACTTCGAC   797

Qy    601  ATTGACATTGTGGCCCGTGGTGAATGACACAGTTGGGACCATGATGAGTGCTATGCAT   660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    798  ATTGACATTGTGGCCCGTGGTGAATGACACAGTTGGGACCATGATGAGTGCTATGCAT   857

Qy    661  GATCAGAACTGCGAGATTGGTCTCATTGTTGTGGCACTGGCAGCAACGCCGCTACACATGGAG   720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    858  GATCAGAACTGCGAGATTGGTCTCATTGTTGTGGCACTGGCAGCAACGCCGCTACACATGGAG   917

Qy    721  GAAATGCGTCATATTGACATGGTGGAGGAGATGAGGGGGCCATGGTGCATCAACATGGAG   780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   1681  TCCATCCCACAGAGGTTATTGCATGGCCACTGGGGAAGAGAGCTCTTCGACCACATTGTCCAG   1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681  TCCATCCCACAGAGGTTATTGCATGGCCACTGGGGAAGAGAGCTCTTCGACCACATTGTCCAG   1740

Qy   1741  TGCATTGCGGACTTCCTTGGAGTACATGGCATGGGCATTGCCTTGGGTTTC   1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741  TGCATTGCGGACTTCCTTGGAGTACATGGCATGGGCATTGCCTTGGGTTTC   1800

Qy   1801  ACATTCTCCTTCCCTTGCCACCAGAGAACAGCCTAGACCAGAGAACAGCCTAGACCAAGAAGTGGACA   1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  ACATTCTCCTTCCCTTGCCACCAGAGAACAGCCTAGACCAAGAAGTGGACA   1860

Qy   1861  AAGGGATTCAAGGCAACTGGCGCTGCCGAGGGTCTGAGGATTGGTCACCTTGCTGGAGGAAGCG   1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861  AAGGGATTCAAGGCACATCTGGCTGCCGAGGGTGGTCACCTTGCTGCTGGAGGAAGCG   1920

Qy   1921  ATTCACCGGCCGGAGGAGTTTGACCTGGAATTGACACACAGTTGGG   1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921  ATTCACCGGCCGGAGGAGTTTGACCTGGAATTGACACACAGTTGGG   1980

Qy   1981  ACTATGATGACTTGTGGCTACGGAAGAGACCCTCACTGTGAAGTTGGCTCATTGTTGCGCACC   2040
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  ACTATGATGACTTGTGGCTACGGAAGAGACCCTCACTGTGAAGTTGGCTCATTGTTGCGCACC   2040

Qy   2041  GGAAGCAACGCCTGCTACATGGAAGAGAGATGCGTAAATGTGGACTGGTGGACGGAGAGGAG   2100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  GGAAGCAACGCCTGCTACATGGAAGAGAGATGCGTAAATGTGGACTGGTGGACGGAGAGGAG   2100

Qy   2101  GGACGGATGTGTGTCAACATGATGAGTGGGGACATTGGGACACATTGCCTGCTGATGAC   2160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  GGACGGATGTGTGTCAACATGATGAGTGGGGACATTGGGACACAATGCCTGGATGAC   2160

Qy   2161  TTGCGGACCGTGTTTGATGTTGCTGTGGATGTGGGATGAGCTTTCTCTCTCAGCCAAAGAGG   2220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  TTGGCGGACCGTGTTGATGTTGCTGTGGATGAGCTTTCTCTCAACCTGGCAAACAGGAGG   2220

Qy   2221  TTCGAGAAGATGATCAGGCGGCATGTACTTGGGCTTCTCTCAAGACAAGGGGA   2280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  TTCGAGAAGATGATCAGGCGGCATGTACTTGGGAGAGATTGTGGCCAACAATTCTCATCGAT   2280

Qy   2281  TTCACGAAAGCGGGGGGGCTGCTCTTCCGAGCCCTCAAGACAAGGGGA   2340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  TTCACGAAAGCGGGGGGGCTGCTCTTCCGAGCCCTCAAGAGCCCTCAAGAGACAAGGGGA   2340

Qy   2341  ATCTCTGAAACTAAGTTCCTGTGTCTCAGATAGAGAGAGGAGACATCATCGGG   2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  ATCTCTGAAACTAAGTTCCTGTGTCTCAGATAGAGAGAGGAGACTGCCTAGCCGTTTACAGGTT   2400

Qy   2401  CGTGCCATCCTGCGCCACCTAGGGCTGGAGAGCATCAACAGCATCATCGTGAAG   2460
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401  CGTGCCATCCTGCGCCACCTAGGGCTGGAGAGCATCAACAGCATCATCGTGAAG   2460

Qy   2461  GAGGTGTGCACTGTGGTTGTTCCCGGCGGCCTGCACAGCTCTGTCTGTGGCCGCGCC   2520
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461  GAGGTGTGCACTGTGGTTGTTCCCGGCGGCCTGCACAGCTCTGTCTGTGGCCGATGGCCGCC   2520

Qy   2521  GTAGTGGACAAGAATAAGACAAACCGGTGGACAGCAGTGACAGTGGGCGTG   2580
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521  GTAGTGGACAAGATAAGACAAGAGAGACCCGGTGGGGCTGGACAACCCCAAAGTGACAGTGGGCGTG   2580

Qy   2581  GACGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATCGATGAGACGGTGAGA   2640
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581  GACGGGACTTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATCGATGAGACGGTGAGA   2640

Qy   2641  GATCTCTGGCCTCCGAAAATGTGACGTGTCCTTCCTGGAATCCGAGGACGGCAGTGGGAAGGGGGA   2700
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641  GATCTCGGCCTCCGAAAATGTGACGTGTCCTTCCTGGAATCCGAGGAGGACGGCAGTGGGAAGGGGA   2700

Qy   2701  GCAGCTCTCATCACTGCCGTGCCGCCGCCCATCCGGGAGGCTGGGCAGAGAGATAG   2754
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701  GCAGCTCTCATCACTGCCGTGCCGCCGCCCATCCGGGAGGCTGGGCAGAGAGATAG   2754
```

```
Db    918 GAAATGGTCATATTGGGGACGGACGGGTACACTCAATGGTGTGGAGGAGATGAGGGCCCATTGTGCATCACAATGGAG 977
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    781 TCGGGAGCCCTTTGGGGACGACGGGTACACTCAATGACATCAATGACACTGGAGGAGGCCCATTGTGACCGAGAG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    978 TGGGGAGCCCTTTGGGGACGACGGGTACACTCAATGACATCAATGACACTGGAGGAGGCCCATTGTGACCGAGAG 1037
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    841 ATCGACAATGGGCTGCGTCGACTGAACCCTGGGAAGCAGCTGTTTGAGAAGATGATTAGCCGGGATG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038 ATCGACAATGGGCTGCGTCGACTGAACCCTGGGAAGCAGCTGTTTGAGAAGATGATTAGCCGGGATG 1097
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021 GATATTGAAGAGATAAGGATGGAATCAGAGAGGAAGGCCTACCCAAATCCTGATGCCCTGGGT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1158 CAAGGGAAAACTCAGCCCAGAAACTCAGCCCAGAAACTCCTTACCACTGTCCTTCGAGACCAAAGATGTCTCG 1217
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    961 CAAGGGAAAACTCAGCCCAGAAACTCAGCCCAGAAACTCCTTACCACTGTCCTTCGAGACCAAAGATGTCTCG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1098 TACATGGGGAGCTGGTCAGGGCTGGCTCATCCTGGTGAAGATGGCCAAGGCGAGCAGCCTGTTGTTC 1157
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901 TACATGGGGGAGCTGTCGCTGGGAAGGCCAGGCTGGTGTTGTTC 960
          |||||||||||||||||||||||||||||||||||||||||||
Db   1081 CTGAATCCAATTGCAGGAGGATTGTGTTGTGGGCCGAATCCAGATTGTGTCCACG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1218 CTGAATCCAATTGCAGGAGGATTAAGGAATCGAGAAGGCCTACCAAATCCTGATGCCCTGGGT 1277
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1278 TTGCGGACCGTGTTTGATGTTGGCTGGTGGGATGAGCTTTCTCTCACACCCTGGCAAACCTGGTGTCCACG 1337
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1141 CGCTGGCCCAGTCTGTGCTGCCGCAGCCAGGATTGTGTGGCCACCACCAGCACCGGAATCAGATTGTGTCCACG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1338 CGCTGGCCGCCGACTCTGTTGCCTCCTCGCCTGGCTGTGTCGGTGTCCGTTGCCGGGAATCAAAGAGAAC 1397
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1398 AAGGGCCGAGGAGCGGACACTTCGCTCCTCCACCATCGGGTGTCGATGGCTTCCGTGCTCTGCTATCAAGAGAAC 1457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AAGGGCCGAGGAGCGGACACTTCGCTCCTCCACCATCGGGTGTCGATGGCTTCCGTGCTCTGCTATCAAGAGAAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1261 CCCCATTTTTGCAAGGCTGTCTCCATAAGGCACGTGAGGACGGTGGTGCGGCCGACTGTGTGATGTC 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1458 CCCCATTTTTGCAAGGCTGTCTCCATAAGGCACGTGAGGACGGTGGTGCGGCCGACTGTGTGATGTC 1517
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1321 CGCTTCCTCCGCTCTCGAGGATGGCACGGCCAAGGGGCTGCTATGGTTGACGGCGGTGGCT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1518 CGCTTCCTCCGCTCTCGAGGATGGCACGGCCAAGGGGCTGCTATGGTTGACGGCGGTGGCT 1577
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1381 TACCGTCTGGCCTGACCAACACCGGCCGCCAGAAGACCCTGGAGTCTCTGAAGCTCAGC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1578 TACCGTCTGGCCTGACCAACACCGGCCGCCAGAAGACCCTGGAGTCTCTGAAGCTCAGC 1637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1441 CACGAGCAGCTTCTGGAGGTTAAGGAGAATCAAAGTGGGAAATGGAGCAGGGTCTCTGAGC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1638 CACGAGCAGCTTCTGGAGGTTACGAGAACAAATGAAAGGTGGGAAATGGAGCAGGGTCTCTGAGC 1697
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1501 AAGGAGACAACCCATGCGGGCTCGCGCCCCTGTGTGAAGATGCTGCCCACTTACGTGTCTGTGCCACTCCA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1698 AAGGAGACAACCCATGCGGGCTCGCGCCCCTGTGTGAAGATGCTGCCCACTTACGTGTCTGTGCCACTCCA 1757
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1561 GATGGCACAGAGAAGAAGAGACTTCTTGGCCTTGACATCTTGGAGGAACAAACTTCCGGGTC 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1758 GATGGCACAGAGAAAAGGAGACTTCTTGGCCTTGACATCTTGGAGGAACAAACTTCCGGGTC 1817
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1621 CTGCCTGGTTCTGTTGCCTAATGCAAGCGAGGGCGTGGGACGCGCATAACACAAGATCTAC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1818 CTGCCTGGTTGTGCCTAATGCAAGCGAGGGCGTGGGACGCGCATAACACAAGATCTAC 1877
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1681 TCCATCCCACAGGAGGTTATGCATGGCACTGGGGAAGAGAGCTCTTCGACCACCACATTGTCCAG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1878 TCCATCCCACAGGAGGTTATGCATGGCACTGGGGAAGAGAGCTCTTCGACCACCACATTGTCCAG 1937
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1741 TGCATTGGCGACTTCCTCGGAGTACATGGGCATGAGACAAAGGGGGCATTGGGTTTC 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1938 TGCATTGGCGACTTCCTGGAGTACATGGGCATGGGACTGGGACAAAGGGGGCATTGGGTTTC 1997
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1801 ACATTCTCCTTCCCTTGCCAGCAGAGAACAGCCTTAGACCAGAGAGCATCCTCCTGCCTTCAAGTGGACA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1998 ACATTCTCCTTCCCTTGCCAGCAGAGAACAGCCTTAGACCAGAGAGCATCCTCCTGCCTTCAAGTGGACA 2057
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1861 AAGGGATTCAAGGCCATCTGGGCTGCGAGGGTGAGGATGTGGTCACCTTGCTGTTGACGGAAGCG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2058 AAGGGATTCAAGGCCATCTGGGCTGCGAGGGTGAGGATGTGGTCACCTTGCTGTTGACGGAAGCG 2117
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1921 ATTCACCGCGAGAGGAGTTTGACCTGCGAGGGTGAGTGTGGTTGCCGGTGGTTGAGTGACACAGTTGG 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2118 ATTCACCGCGAGAGGAGTTTGACCTGCGAGGGTGAGTGTGGTTGCCGGTGGTTGAGTGACACAGTTGG 2177
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1981 ACTATGATGACTTGTTGCTACGAGCGAAGACCCTCACTGTGAAGTTGGCCTTCAGTTGGCACC 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2178 ACTATGATGACTTGTTGCTACGAGCGAAGACCCTCACTGTGAAGTTGGCCTTCAGTTGGCACC 2237
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2041 GGAGCAACGCGCTGCTACATGGAAGAGATAGCGTAAATGTGGACTTGGGACTGGGA 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2238 GGAGCAACGCGCTGCTACATGGAAGAGATAGCGTAAATGTGGACTTGGGACTGGGA 2297
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2101 GGACGGAATGGTGTTGTCAACATGGAGTAGGAGTGGGGAGCACATTTGGGGACAATGGCTGCCTGGGATGAC 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2298 GGACGGAATGGTGTTGTCAACATGGAGTAGGAGTGGGGAGCACATTTGGGGACAATGGCTGCCTGGGATGAC 2357
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2161 TTGCGGACCGTGTTTGATGTTTGCTGTGGATGAGCTTTCTCTCACACCCTGGCAAACACAGAGG 2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2358 TTGCGGACCGTGTTTGATGTTTGCTGTGGATGAGCTTTCTCTCACACCCTGGCAAACACAGAGG 2417
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2221 TTCGAGAAGAGATGATCAGCGAGGGCGTGTACTTGGGGAGAGATTGTGTGCCGGAACAATTCTCATCGAT 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2418 TTCGAGAAGAGATGATCAGCGAGGGCGTGTACTTGGGGAGAGATTGTGTGCCGGAACAATTCTCATCGAT 2477
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2281 TTCGAGAAGAACTAAGTTCCGTGTCTCAGATAAGAGAGCGGCCATCTCAAGACACAAGGGGA 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2478 TTCACGAACGGGGGGGGGCTGCCGCCATCTCAGATAAGAGAGCGGCCATCTCAAGACACAAGGGGA 2537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2341 ATCTTTGAAACTAAGTTCCGTGTCTCAGATAAGAGAGCGGCCGCTTCAGAAGACGCCTAGCCCTCCTACAGGTT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2538 ATCTTTGAAACTAAGTTCCGTGTCTCAGATAAGAGAGCGGCCGCTTCAGAAGACGCCTAGCCCTCCTACAGGTT 2597
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2401 CGTGCCATCCTGCCCACCTAGGCGCTGGACTGGACGGACGTGCTGTGTCGATGACAGCAGCGTGAAG 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2598 CGTGCCATCCTGCCCACCTAGGCGCTGGACTGGACGGACGTGCTGTGTCGATGACAGCAGCGTGAAG 2657
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2461 GAGGTGTTGCACTGTGTGGTTGCCCGCGGCCACCTAGGCCATCATCGTGGCGCC 2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2658 GAGGTGTTGCACTGTGTGGTTGCCCGCGGCCACCTAGGCCATCATCGTGGCGCC 2717
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2521 GTAGTGGACAAGAATAAGAGAGAACCGTGGGGCGTGGGACAACACCCAAAGTGACAGTGGGCGTG 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2718 GTAGTGGACAAGAATAAGAGAGAACCGTGGGGCGTGGGACAACACCCAAAGTGACAGTGGGCGTG 2777
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2581 GACGGGACTCTGTATAAGCTTCATCCTCACTTTTGCCAAGTCATGAGACGGGTGAGA 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2778 GACGGGACTGTTATAAGCTTCATCCTCACTTTTGCCAAGTCATGAGACGGGTGAGA 2837
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2641 GATCTGGCTCGCCGAAAATGTGACGTGTCCTTCCTTCTTGGAATCCGAGGACGGGAGGGGA 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2838 GATCTGGCTCGCCGAAAATGTGACGTGTCCTTCCTTCTTGGAATCCGAGGACGGGAGGGGA 2897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2701 GCAGCTCTCATCACTGCCGTGGCCTGGGCCGCCGCATCCGGGGGAGGCTGGGCGAGAGATAG 2754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2898 GCAGCTCTCATCACTGCCGTGGCCTGGGCCGCCGCATCCGGGCCATCCGGGGGAGGCTGGGCGAGAGATAG 2951
```

RESULT 3
AR070139

```
REFERENCE    1 (bases 1 to 3635)
  AUTHORS    Newgard,C.B., Han,H.-P., Becker,T.C. and Wilson,J.E.
  TITLE      Methods and compositions for inhibiting hexokinase
  JOURNAL    Patent: US 5891717-A 15 06-APR-1999;
FEATURES             Location/Qualifiers
     source          1..3635
                     /organism="unknown"
BASE COUNT      816 a    875 c    1098 g    846 t
ORIGIN

Query Match     92.6%;  Score 2550;  DB 6;  Length 3635;
Best Local Similarity  99.9%;  Pred. No. 0;
Matches  2750;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;
```

Alignment (Query vs Database), positions:

```
Qy     1 ...    60        Db   198 ...   257
Qy    61 ...   120        Db   258 ...   317
Qy   121 ...   180        Db   318 ...   377
Qy   181 ...   240        Db   378 ...   437
Qy   241 ...   300        Db   438 ...   497
Qy   301 ...   360        Db   498 ...   557
Qy   361 ...   420        Db   558 ...   617
Qy   421 ...   480        Db   618 ...   677
Qy   481 ...   540        Db   678 ...   737
Qy   541 ...   600        Db   738 ...   797
Qy   601 ...   660        Db   798 ...   857
Qy   661 ...   720        Db   858 ...   917
Qy   721 ...   780        Db   918 ...   977
Qy   781 ...   840        Db   978 ...  1037
Qy   841 ...   900        Db  1038 ...  1097
Qy   901 ...   960        Db  1098 ...  1157
Qy   961 ...  1020        Db  1158 ...  1217
Qy  1021 ...  1080        Db  1218 ...  1277
Qy  1081 ...  1140        Db  1278 ...  1337
Qy  1141 ...  1200        Db  1338 ...  1397
Qy  1201 ...  1260        Db  1398 ...  1457
Qy  1261 ...  1320        Db  1458 ...  1517
Qy  1321 ...  1380        Db  1518 ...  1577
Qy  1381 ...  1440        Db  1578 ...  1637
Qy  1441 ...  1500        Db  1638 ...  1697
Qy  1501 ...  1560        Db  1698 ...  1757
Qy  1561 ...  1620        Db  1758 ...  1817
Qy  1621 ...  1680        Db  1818 ...  1877
Qy  1681 ...  1740        Db  1878 ...  1937
Qy  1741 ...  1800        Db  1938 ...  1997
Qy  1801 ...  1860        Db  1998 ...  2057
Qy  1861 ...  1920        Db  2058 ...  2117
Qy  1921 ...  1980        Db  2118 ...  2177
```

```
RESULT 4
LOCUS       RATHKII      3635 bp    mRNA        linear   ROD 27-APR-1993
DEFINITION  Rat hexokinase type II (HKII) mRNA, complete cds.
ACCESSION   M68971
VERSION     M68971.1  GI:204612
KEYWORDS    hexokinase type II.
SOURCE      Rat (strain C.R. sd) skeletal (soleus) muscle, cDNA to mRNA.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 3635)
  AUTHORS   Thelen,A.P. and Wilson,J.E.
  TITLE     Complete amino acid sequence of the type II isozyme of rat
            hexokinase, deduced from the cloned cDNA: comparison with a
            hexokinase from novikoff ascites tumor
  JOURNAL   Arch. Biochem. Biophys. 286 (2), 645-651 (1991)
   MEDLINE   91378366
   PUBMED    1897984

FEATURES             Location/Qualifiers
     source          1..3635
                     /organism="Rattus norvegicus"
                     /strain="c. R. sd"
                     /db_xref="taxon:10116"
                     /tissue_type="soleus muscle"
                     /tissue_lib="lambda gt10"
     gene            1..3635
                     /gene="HKII"
     CDS             198..2951
                     /gene="HKII"
                     /EC_number="2.7.1.1"
                     /codon_start=1
                     /product="hexokinase type II"
                     /protein_id="AAA41333.1"
                     /db_xref="GI:204613"
                     /translation="MIASHMIACLFTELNQNQVKDQFLYHMRLSDETLLEISRRFR
                     KEMEKGLGATTHPTAAVKMLPTFVRSTPDGTEHGEFLADLDGTNFRVLRVRTDNGL
                     QRVEMENQIYAIPEDIMRGSGTQLFDHIAECLANFMDKLQIKEKRLPLGFTFSFPCHQ
                     TKLDESFLVSWTKGFKSSGVEGRDVVDLIRKAIQRRGDFDIDIVAVVNDTVGTMMYCG
                     YDDQNCEIGLIVGTGSNACYMEEMRHIDMVEGDEGRMCINMEWGAFGDGTLNDIRTE
                     FDREIDMGSLNPGKQLFEKMISGMYLGELVRLILVKMAKEELLFGGKLSPELLTTGSF
                     ETKDVSDIEEDKDGIEKAYQILMRLGLNPLQEDCVATHRICQIVSTRSASLCAATLAA
                     VLWRIKENKGEERLRSTIGVDGSVYKKHPHFAKRLHKAVRRLVPDCDVRFLRSEDGSG
                     KGAAMVTAVAYRLADQHRARQKTLESLKLSHEQLLEVKRRMKVEMEQGLSKETHAVAP
                     VKMLPTYVCANPDGTEKGDFLALDLGGTNFRVLLVRVRNGKRRGVEMHNKIYSIPQEV
                     MHGTGEELFDHIVQCIADFLEYMGMKGVSLPLGFTFSFPCQQNSLDQSILLKWTKGFK
                     ASGCEGEDVVTLLKEAIHRREEFDLDVVAVVNDTVGTMMTCGFEDPHCEVGLIVGTGS
                     NACYMEEMRNVELVDGEEGRMCVNMEWGAFGDNGCLDDLRTVFDVDAVADELSLNPGKQR
                     FEKMISGMYLGEIVRNILIDFTKRGLLFRGRISERLKTRGIFETFKTLSQIESDCLALL
                     QVRAILRHLGLESTCDDSIIVKEVCTVVARRAAQLCGAGMAAVVDKIRENRGLDNLKV
                     TVGVDGTLYKLHPHFAKVMHETVRDLAPKCDVSFLESEDGSGKGAALITAVACRIREA
                     GQR"

BASE COUNT      816 a    875 c   1098 g    846 t
ORIGIN

Query Match              92.6%;   Score 2550;   DB 10;   Length 3635;
Best Local Similarity    99.9%;                           Pred. No. 0;
Matches 2750;   Conservative   0;   Mismatches   4;   Indels   0;   Gaps   0;
```

```
Qy     1 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACCGAGCTCAACCAAAACCAAGTGCAG   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACCGAGCTCAACCAAAACCAAGTGCAG  257

Qy    61 AAGGTTGACCAAATTTCTCTACCACACATGCGTCTCTCAGATGAGACCCTTCTGGAGATTTCT  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258 AAGGTTGACCAAATTTCTCTACCACACATGCGTCTCTCAGATGAGACCCTTCTGGAGATTTCT  317

Qy   121 AGGCGGTTCCGGAAGGAGAGATGGAGAAAAGGGCTAGGAGCTACCACGACCCTTACAGCAGCT  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318 AGGCGGTTCCGGAAGGAGAGATGGAGAAAAGGGCTAGGAGCTACCACGACCCTTACAGCAGCT  377

Qy   181 GTGAAAGTGTTGCCTACCTTTGTGAGGTCAACTCCGGATGGGACAGAGACAGAGAGTTC  240
         ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   378 GTGAAAAATGTTGCCTACCTTTGTGAGGTCAACTCCGGATGGGACAGAGACAGAGAGTTC  437

Qy   241 CTGGCTCTTGGATCTTGGAGGAACCAACTTCCGTGTTGCTGCTCCCGAGTGGGACACAAT  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   438 CTGGCTCTTGGATCTTGGAGGAACCAACTTCCGTGTTGCTGCTCCCGAGTGGGACACAAT  497

Qy   301 GGCCTTCCAGAGAGTGGAGATGGAGAACCAGATCTACGCCATCATCCTTGAGGACATCATGCGG  360
         |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   498 GGCCTTCCAGAGAGTGGAGATGGAGAACCAGATCTACGCCATCCTTGAGGACATCATGCGG  557

Qy   361 GGCAGTGGAACCAGCTGTTTGACCACATCGCCGAATGCCTGGCCGACACCATGGGGGAGTTC  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   558 GGCAGTGGAACCAGCTGTTTGACCACATCGCCGAATGCCTGGCCGACACCATGGGGGAGTTC  617

Qy   421 CTACAAATCANAAGAGAAGAAAGAGCTCCCTCTGGGTTTCACCTTCTCTTTCCCTGCCACCAG  480
         |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   618 CTACAAATCAAAAGAGAAGAAAGAGCTCCCTCTGGGTTTCACCTTCTCTTTCCCTGCCACCAG  677

Qy   481 ACAAAACTGGATGAGGAGTTTTTGGTTCTCGTGGACTAAGGGGTTCAAGTCAGTGGCGTG  540
```

```
Qy  1981 ACTATGATGACTTGTGGCTACGAGAGACCCTCACTGTGAAGTTGGCCTCATTGTTGGCACC  2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2178 ACTATGATGACTTGTGGCTACGAGAGACCCTCACTGTGAAGTTGGCCTCATTGTTGGCACC  2237

Qy  2041 GGAAGCAACGCCTGCTCACATGGAAGAGATGCGTAATGTGGAGCTGGTGGACGGAGGAGGAG  2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2238 GGAAGCAACGCCTGCTCACATGGAAGAGATGCGTAATGTGGAGCTGGTGGACGGAGGAGGAG  2297

Qy  2101 GGACGGATGTGTGTCAACATGAGTGGGGAGCAATGGGCTGCCTGACCGATGAC  2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2298 GGACGGATGTGTGTCAACATGAGTGGGGAGCAATGGGCTGCCTGACCGATGAC  2357

Qy  2161 TTGCGGACCGTGTTTGATCGTGTTGTGTGTGGATGAGGAGCAATTTGGGGACAATGGCTGCC  2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2358 TTGCGGACCGTGTTTGATCGTGTTGTGTGTGGATGAGGAGCAATTTGGGGACAATGGCTGCC  2417

Qy  2221 TTCGAGAAGATGATCAGCGGCATGTACTTGGGAGAGATTGTGCGCCAACATTCTTCATCGAT  2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2418 TTCGAGAAGATGATCAGCGGCATGTACTTGGGAGAGATTGTGCGCCAACATTCTTCATCGAT  2477

Qy  2281 TTCACGAAGCGGGGGCTGCTCTTCCGAGGCCGCATCTCAGAGCGCCTCAAGACAAGGGGA  2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2478 TTCACGAAGCGGGGGCTGCTCTTCCGAGGCCGCATCTCAGAGCGCCTCAAGACAAGGGGA  2537

Qy  2341 ATCTTCGAAACTAAGTTCCTGTCTCTCAGATAGAGAGCGACTAGCCCTGCTACAGGTT  2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2538 ATCTTCGAAACTAAGTTCCTGTCTCTCAGATAGAGAGCGACTAGCCCTGCTACAGGTT  2597

Qy  2401 CGTGCCATCCTGCGCCACCTAGGGCTGGAGAGCACGTGCGATGACATCATCGTGAAG  2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2598 CGTGCCATCCTGCGCCACCTAGGGCTGGAGAGCACGTGCGATGACATCATCGTGAAG  2657

Qy  2461 GAGGTGTGTCACTGTGGTTGGTTGCCCCGGCCGCCGGGGCCGCTCGCACAGCGCTCTCGTGCCCC  2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2658 GAGGTGTGTCACTGTGGTTGGTTGCCCCGGCCGCCGGGGCCGCTCGCACAGCGCTCTCGTGCCCC  2717

Qy  2521 GTAGTGGACAAGAAGATAAGAGAGAACCGTGGGCGGTGGACACACCCCAAAGTGACAGTGGGCCGTG  2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2718 GTAGTGGACAAGAAGATAAGAGAGAACCGTGGGCGGTGGACACACCCCAAAGTGACAGTGGGCCGTG  2777

Qy  2581 GACGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATGCAGAACACGGTGAGA  2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2778 GACGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATGCAGAACACGGTGAGA  2837

Qy  2641 GATCTGGCTCCCAAAATGTGACGTGTGTCTTCCTTCCTCACTTGTCATCGAGAGATTTCT  2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2838 GATCTGGCTCCCAAAATGTGACGTGTGTCTTCCTTCCTCACTTGTCATCGAGAGATTTCT  2897

Qy  2701 GCAGCTCTCATCACTGTGGTTGCGTGCCGTGCCCGTGCCGGAGGCTGGGCACAGAGATAG  2754
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2898 GCAGCTCTCATCACTGTGGTTGCGTGCCGTGCCCGTGCCGGAGGCTGGGCACAGAGATAG  2951
```

Db 678 ACAAAACTGGATGAGAGAGTTTTTGGGTCTCGTGGACTAAGGGGTTCAAGTCCAGTGGCGTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 541 GAAGGCAGAGATGTGGTGGACCTGATCCGGAAGGTTTATCCACGGGACGACTTTGAC 600

Db 738 GAAGGCAGAGATGTGGTGGACCTGATCCGGAAGGCTATCCACGGGAGGGGACTTTGAC 797
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 601 ATTGACAATTGTGGCCGTGGTGAATGACACAGTTGGGACTTGTGGCTATGAT 660

Db 798 ATTGACAATTGTGGCCGTGGTGAATGACACAGTTGGGACCATGATGACTTGTGGCTATGAT 857
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 661 GATCAGAACTGCCGAGATTGGTCTCATTGTGGGGCACTGGCAGCAGCGCCTTGCTACATGGAG 720

Db 858 GATCAGAACTGGCGTCGTTGAACCTGTGGGCACTGGCAGCAGCGCCTGCTACATGGAG 917
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 721 GAAATGCGTCATATTGACATGGTGGAGGGAAGGGCATGTGCATCAACATGGGA 780

Db 918 GAAATGCGTCATATTGACATGGTGGAGGGGAGGGCCATGTGCATCAACATGGGA 977
||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 781 TGGGGACCCTTTGGGACGGACGGTACACTCAATGACACAACCGACTTTGACCGAGAG 840

Db 978 TGGGGACCCTTTGGGACGGACGGTACACTCAATGACACCGACTTTGACCGAGAG 1037
|||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 841 ATCGACAATGGGCTCGCTGCTGAACCCTGGGAGAGCAGCTGTTTGAGAGGATGA 900

Db 1038 ATCGACAATGGGCTCGCTGAACCCTGGGAGAGCAGCTGTTTGAGAAGATGAATTACGGGGATG 1097
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 901 TACATGGGAGAGCTGGTCAGGGTCATCCTGGGTGAAGAC 960

Db 1098 TACATGGGAGAGCTGGTCAGGGTCATCCTGGGTGAAGAC 1157
|||||||||||||||||||||||||||||||||||||||
Qy 961 CAAGGGAAAACTCAGCCAGAACTCCTTACCACTGGCCTCCTTGTTGAGAATGTCTCG 1020

Db 1158 CAAGGGAAAACTCAGCCCAGAACTCCTTACCACTGGCCTCCTTGTTGAGAATGTCTCG 1217
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1021 GATATTGAAGAGAGGGAATAAGGATGGAATGGGAGAGGGGCCTACCAAAGTGTCCACG 1080

Db 1218 GATATTGAAGAGGATAAGGATGGAATGGGAGAGGGGGCCTACCAAAGTGTCCACG 1277
||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1081 CTGAATCCATTGCAGGATAAGGCTTCCACTGGCCTTACCAAATCCTGGGT 1140

Db 1278 CTGAATCCATTGCAGGATAAGGCTTCCACTGGCCTTACCAAATCCTGATGTGTTGTTC 1337
|||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1141 CGCTCGGCGCAGTCTGTGCGGCGCAGCCACCCTGGCGCTGGGCGGAATCAACAGAAC 1200

Db 1338 CGCTCGGCGCAGTCTGTGCGGCGCAGCCACCCTGGCGCTGGGCGGAATCAACAGAAC 1397
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1201 AAGGGCGGAGGGAGCGGACTTCGCTCCACCATCGGTGTCGATGGCGTCTCCGTTCTACAAGAACAT 1260

Db 1398 AAGGGCGGAGGGAGCGGACTTCGCTCCACCATCGGTGTCGATGGCGTCTCCGTTCTACAAGAACAT 1457
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1261 CCCAATTTGCCAAGCGTTCTCCATAAGGCACTGAGGCGGCTGGTCGGTGCCCGACTGTGATGTC 1320

Db 1458 CCCAATTTGCCAAGCGTTCCATAAGGCACTGAGGCGGCTGGTCGGTGCCCGACTGTGATGTC 1547
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1321 CGCTTCCTCCGTCTGAGGATGGCAGGGGGCTGCTATGGTCACGGCGGTGGCT 1380

Db 1518 CGCTTCCTCCGTCTGAGGATGGCAGGGGGCTGCTATGGTCACGGCGGTGGCT 1577
|||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1381 TACCGTCTGGGCTGACCAACACCGGGCCCCGCAGAGACACCCTGGGAGTCTCTGAAGCTGAGC 1440

Db 1578 TACCGTCTGGGCTGACCAACACCGGGCCCCGCAGAGACACCCTGGGAGTCTCTGAAGCTGAGC 1637
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1441 CACGAGCAGCAGCTTCTGGAGGTTAAGAGAGAATGAAGGTGGAAATGGAGCAGGGTCTGAGC 1500

Db 1638 CACGAGCAGCAGCTTCTGAGGGTTAAGAGAGAATGAAGGTGGAAATGGAGCAGGGTCTGAGC 1697
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1501 AAGGAGAGCACATGCGGTCGCCCCGTGTGAAAGATGCTTACGTGTGTGTGCCACTCCA 1560

Db 1698 AAGGAGAGCACATGCGGTCGCCCCGTGTGAAAGATGCTTACGTGTGTGTGCCACTCCA 1757
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1561 GATGGGCACAGAGAAAAGGAGAGACTTCTTGGGATCTTGGAGGACAAAACTTCCGGGTC 1620

Db 1758 GATGGGCACAGAGAAAAGGAGAGACTTCTTGGGATCTTGGAGGACAAAACTTCCGGGTC 1817

Db 1758 GATGGGCACAGAGAAAAGGAGAGACTTCTTGGCCTTGGATCTTGGAGGAACAAAACTTCCGGGTC 1817
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1621 CTGCTGGTGCTGGTGTCGTAATGCAAGGCATGGCAGGATCATAACAGATCTAC 1680

Db 1818 CTGCTGGTGCGCGTGTTGCGTAATGGCAAGGGCGGAGGGATGCATAACAAGATCTAC 1877
|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1681 TCCCATCCCACAGGAGAGGTTTATTGCCATGGCACTGGGGAAGGACTCTTCGACACACAACATTGTCCAG 1740

Db 1878 TCCCATCCCACAGGAGAGGTTTATTGCCATGGCACTGGGGAAGGACTCTTCGACACACAACATTGTCCAG 1937
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1741 TGCATTGCCGGACTTCCTGGAGTACATGGGGATCATGGGGGCACTGAAGGGGCCATGGGGGCCTTTGGGTTTC 1800

Db 1938 TGCATTGCCGGACTTCCTGGAGTACATGGGGATCATGGGGGCACTGAAGGGGCCATGGGGGCCTTTGGGTTTC 1997
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1801 ACATTCTCCTTCCCTTGCCAGCAGAACAGCCTAGACCAGCAGCCATCCTTCAACTCAAGTGGA 1860

Db 1998 ACATTCTCCTTCCCTTGCCAGCAGAACAGCCTAGACCAGCAGCCATCCTTCAACTCAAGTGGA 2057
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1861 AAGGGATTCAAAGGCACATCTGGCTGGAGGCATTCTGGCCTACAAGGTCAAGTTGGACA 1920

Db 2058 AAGGGATTCAAAGGCCATCTGGCCTGCCAGGGTTATGCATTGCCACTGGGGAAGAGCTCTACAAGGTCAAGTTGGACA 2117
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1921 ATTCACCGGCGGAGGAGGTTTGACCTGGATGTGCTGGAGTTGGATGACATGGGATCAGTTGGG 1980

Db 2118 ATTCACCGGCGGAGGAGGTTTGACCTGGATGTGCTGGAGTTGGATGACATCAGTTGGG 2177
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1981 ACTATGATGACTTGTGGCTACGAAGAGAACAGCCTAGAGAACCCTCACTCTGTGAAGTTGGCCACC 2040

Db 2178 ACTATGATGACTTGTGGCTACGAAGAGAACAGCCTAGAGAACCCTCACTCTGTGAAGTTGGCCACC 2237
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2041 GGAGCAGCAAGCGCCTCGTACGAAGGAGAGATGCGTAAGTGCCAGGAGGAG 2100

Db 2238 GGAAGCAACAGCCCTCGTACGAAGGAGAGATCGTAAGTGGGGGGTGGAGAGGAG 2297
||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2101 GGACGGATGTGTGTCAACATGGGAGTGGGGACAAATGGCTTGCCTGGATGAC 2160

Db 2298 GGACGGATGTGTGTCAACATGGGAGTGGGGACAAATGGCTTGCCTGGATGAC 2357
|||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2161 TTCAGAGAAGATGATCAGGGCAGTGTTACTGGGCGCCATCACCCTGGCCAAACGAGG 2220

Db 2358 TTGCGAGAAGATGATCAGCGGCAGTTACTTGGGGGATTGTGGCGCCAATCTCCATCGAT 2417
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2221 TTCACGAACCGGGGGGCTGCTCTTCCGAGGGCGCATCCTCAGAGACAAGGGGGA 2280

Db 2418 TTCACGAACCGGGGGGCTGCTCTTCCGAGGGCGCATCTCAGAGACAAGGGGGA 2477
||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2281 ATCTCTGAAACTAAGTTCCTGTCTCAGATAGGAGAGACGGGGGTT 2340

Db 2478 ATCTTTGAAACTAAGTTCCTGTCTCGAGATAGAGAGACGCCTGACCAGGGTT 2537
||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2341 GACGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGTCATGCAGTGAGACGGGTGAGA 2400

Db 2538 GACGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATGCAGAGACGGGTGAGA 2597
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2401 CGTGCCCATCCTGCGCCACCTAGGGGCTGGGAGGAGCAGCAGCATCATCGTGAAG 2460

Db 2598 CGTGCCCATCCTGCGCCACCTAGGGGCTGGGAGGAGCAGCAGCATCATCGTGAAG 2657
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2461 GAGGTGTGCACTGTTGGTTGCCGCGGCGGCTGGAGGAGCACAGCACTGGCCGCC 2520

Db 2658 GAGGTGTGCACTGTTGGTTGCCCGCGGCGGCTGGAGGAGCACAGCACTGGCCGCGCC 2717
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2521 GTAGTGGACAAGATAAGGAGGAACCGTGGGCGGGCGGGGGCGCTGGGGGGGGTG 2580

Db 2718 GTAGTGGACAAGATAAGGAGGAACCGTGGGCGGGCGGGGGCGCTGGGGGGGCGG 2777
||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2581 GACGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATGCAGAGACGGGTGAGA 2640

Db 2778 GACGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATGCAGAGACGGGTGAGA 2837
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2641 AAGGAGAGCACATGCGGTCGCCCCGTGTGAAAGATGCTTACGTGTGTGTGCCACTCCA 2700

Db 2838 GATCTGGCTCCGAAATGTGACGTGTCCTTCCTTGGAAATCCGAGGACGGACGACGGCAGTGGGAAGGGA 2897

```
Qy  2701 GCACCTTCTCATCACTGCGTGCCGTGCCGTGCCGTGGCGTCCGGAGGGCTGGGCGCAGAGATAG 2754
         ||||||||||||||||||||||| || ||||||||||||    |||||||| |||||||||||||
Db  2898 GCACCTCTCATCACTGCCGTGCCGTGCCGTGCCGGCCATCCGGAGGGCAGGCTGGGCAGAGATAG 2951
```

RESULT 5

LOCUS   RATHKIINAH

LOCUS   RATHKIINAH   2236 bp   mRNA   linear   ROD 27-APR-1993
DEFINITION Rat hexokinase II (HKII) mRNA, 3' end.
ACCESSION   M68972
VERSION   M68972.1 GI:204514
KEYWORDS   hexokinase type II.
SOURCE   Rat Novikoff ascites hepatoma, cDNA to mRNA.
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 2236)
  AUTHORS   Thelen,A.P. and Wilson,J.E.
  TITLE   Complete amino acid sequence of the type II isozyme of rat
            hexokinase, deduced from the cloned cDNA: comparison with a
            hexokinase from novikoff ascites tumor
  JOURNAL   Arch. Biochem. Biophys. 286 (2), 645-651 (1991)
  MEDLINE   91378366
   PUBMED   1897984
FEATURES   Location/Qualifiers
     source   1..2236
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /tissue_type="novikoff ascites hepatoma"
     gene   1..2236
                     /gene="HKII"
     CDS   <1..1552
                     /gene="HKII"
                     /codon_start=2
                     /product="hexokinase type II"
                     /protein_id="AAA41334.1"
                     /db_xref="GI:204615"
                     /translation="GEERLRSTIGVDGSVYKKHPHFAXRLHKAVRRLYPDCDVRFLRS
                     EDGSGKGAAMVTAVAYRLADQHRAROKTLESLKLSHEQLLEVKRRMKVEMEOGLSKET
                     HAVAPVKMLPTYVCATPDGTETGDFLALDLGGTNFRVLLVRVRNGKGVEMHNKIYS
                     IPQEVMHGTGEELFDHIVQCIADFLEYMGMKGVSLPLGFTFSFPCQONSLDQSILLKW
                     TKGFKASGCEGEDVTTLKEAIHRREEFDLDVVAVVNDTVGTMMTCGYEDPHCEVGLI
                     VGTGSNACYMEEMRNVELVDGEGRMCVNMEWGAFGDNGCLDDLRTVFDVAVDELSLN
                     PGKORFEKMISGMYLGEIVRNILIDFTKRGLLFRGRISERLKTRGIFETKFLSQIESD
                     CLALLQVRAILRIHLGLESTCDDSIIVKEVCTVARRAAQLCGAGMAAVVDKIRENRGL
                     DNLKVTVGVDGTLYKLHPHFAKVMHETVRDLAPKCDVSFLESEDGSGKGAALITAVAC
                     RIREAGOR"
BASE COUNT   485 a   524 c   680 g   547 t
ORIGIN
```

Query Match     52.7%;  Score 1450;  DB 10;  Length 2236;
Best Local Similarity   99.9%;  Pred. No. 0;
Matches 1550;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

```
Qy  1203 GGGGGAGGAGCGGACTTCGGTCGCTCCACCATCCGGTGTTGGATGGCTTCCCTCTTCACAAGAAAACATCC 1262
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GGGGGAGGAGCGGACTTCGGTCGCTCCACCATCCGGTGTGCTCCGTGCTCCGTCTTCCGTCTTCTACAAGAAAACATCC 60

Qy  1263 CCAATTTGCCAAGGCTCTCCATAAGGCAGCAGTGAGGAGGCTGGTCCGGACTGTGATGTCCG 1322
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CCAATTTTCCCAAGGCTCTCCATAAGGCAGCAGTGAGGAGGCTGGTCCGGACTGTGATGTCCG 120

Qy  1323 CTTCCTCTCCGCTCTGAGGATGGCAAGGCAGCCAAGGGGGCTGCTCGTATTGGTGACGGCGGTGCCTTA 1382
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTTCCTCCCCGCTCTGAGGATGGCAAGGCAGCCAAGGGGGCTGCTCGTATTGGTGACGGCGGTGGCTTA 180

Qy  1383 CCGTTCTGGCTGACCAACACCGGCCCGCCCCAGAAGAGACCCTGGAGTCTCTCAAGCTGAGCCA 1442
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGTTCTGGCTGACCAACACCGGCCCGCCCCAGAAGAGACCCTGGAGTCTTCTGAAGCTGAGCCA 240

Qy  1443 CGAGCAGCTTCTGGAGGTTAAGAGAAGAATGAAGTGGAAATGAAGGAGAGGAGGTTCTGAGCAA 1502
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   241 CGAGCAGCTTCTGGAGGTTAAGAGAAGAATGAAGTGGAAATGAAGGAGAGGAGGTTCTGAGCAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1503 GGAGAGGCATGCGGTCCCGTGTGAAGATGCTGCCCTGTTGGATCTTGGGATGGCCTTGGGGTTCCT 1562

Db   301 GGAGAGGCCATGCGGTCCCGTGTGAAGATGCTGCCCTGTTGGATCTTGGGATGGCCTTGGGGTTCCT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1563 TGGCACAGAGAAAAGGAGGACTTCTTTGGCCTTGGGGATCTTGGAGATGCTGCCACTCCAGA 1622

Db   361 TGGCACAGAGAAAAGGAGGACTTCTTTGGCCTTGGGGATCTTGGAGATGCTGCCACTCCAGA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1623 GCTGGTTCGTGTGCGTAATGCAAGCGGAGCGGAGGGCGTGGAGATGCATAACAAGATCTACTC 1682

Db   421 GCTGGTTCGTGTGCGTAATGCAAGCGGAGCGGAGGGCGTGGAGATGCATAACAAGATCTACTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1683 CATCCCACAGGAGGTTATGCATGGCACTGGGGAAGAGCTCTTCGACCACATTGTCCAGTG 1742

Db   481 CATCCCACAGGAGGTTATGCATGGCACTGGGGAAGAGCTCTTCGACCACATTGTCCAGTG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1743 CATTGCGGACTTCCTGGAGTACATGGGCATGAAGGGCGTGTCCCTGCCTTTGGGTTTCAC 1802

Db   541 CATTGCGGACTTCCTGGAGTACATGGGCATGAAGGGCGTGTCCCTGCCTTTGGGTTTCAC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1803 ATTCTCCTTCCCTTGCCAGCAGAACAGCCTAGACCAGAGCATCCTCCTCAAGTGGACAAA 1862

Db   601 ATTCTCCTTCCCTTGCCAGCAGAACAGCCTAGACCAGAGCATCCTCCTCAAGTGGACAAA 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1863 GGGGATTCAAGGCAGCATCTGGCCTGCCAGGGGTGAGGAGTGTGGTCACCTTGCTGAAGGAAGCCGAT 1922

Db   661 GGGATTCAAGGCCATCTGGCCTGCCAGGGGTGAGGAGTGTGGTCACCTTGCTGAAGGAAGCGGAT 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1923 TCACCCGGCGAGAGGAGAGTTTGACCTGGATGTGGTTGCCGTGGTGAATGACACAGTTGGGGAC 1982

Db   721 TCACCCGGCGAGAGGAGAGTTTGACCTGGATGTGGTTGCCGTGGTGAATGACACAGTTGGGGAC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1983 TATGATGGACTTGTGGTCTGGCTGGACGTTGGCCTTCATTGTTGGTGCAGTTGCGCACCGG 2042

Db   781 TATGATGGACTTGTGGTCTGGCTGGACGTTGGCCTTCATTGTTGGTGCAGTTGCGCACCGG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2043 AAGCAACGCCTGCTACATGGAGATGGCGTAATGTGGAGCTGGTGGTGGACGAGGAGGAGG 2102

Db   841 AAGCAACGCCTGCTACATGGAGATGGCGTAATGTGGAGCTGGTGGTGGACGAGGAGGAGG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2103 ACGGGATGTGTGTCAACATGGAGTGGGGAGCATTTGGGGACAATGGCTGCCTGGATGACTT 2162

Db   901 ACGGGATGTGTGTCAACATGGAGTGGGGAGCATTTGGGGACAATGGCTGCCTGGATGACTT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2163 GCGGGACCGTGTTTGATGTTGCTGTGGATGAGCTTTCTCTCAACCCTGGCAAACAGAGGTT 2222

Db   961 GCGGGACCGTGTTTGATGTTGCTGTGGATGAGCTTTCTCTCAACCCTGGCAAACAGAGGTT 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2223 CGAGAAGATGATCAGCGGGATGTACTTGGGAGAGATAGTGCGGAACATTCTCATCGATTT 2282

Db  1021 CGAGAAGATGATCAGCGGGATGTACTTGGGAGAGATAGTGCGGAACATTCTCATCGATTT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2283 CACGAAACGGGGGCTGCTGTTCCGAGGCCGCATCTCAGAGCGCCTCAAGACAAGGGGAAT 2342

Db  1081 CACGAAACGCGGGCTGCTGTTCCGAGGCCGCATCTCAGAGCGCCTCAAGACAAGGGGAAT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2343 CTTTGAAACTAAGTTCCTGTCTCGTATAGAGAGCGACTGCCTGGCCCTACTCCAAGTTCG 2402

Db  1141 CTTTGAAACTAAGTTCCTGTCTCGTATAGAGAGCGACTGCCTGGCCCTACTCCAAGTTCG 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2403 TGCCATCCTGCGGCCACCTAGGGCTGAGAGCACGTGCGATGACAGCATCATCGTGAAGGA 2462

Db  1201 TGCCATCCTGCGGCCACCTAGGGCTGAGAGCACGTGCGATGACAGCATCATCGTGAAGGA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2463 GGTGTGTCACTGTGGTTCCCGGCGCTGCACGCGCAGGGCTGCCCAGCTCTGTGGGGCCGT 2522

Db  1261 GGTGTGTCACTGTGGTTGCCCGGCGCTGCACGCGCAGGGCTGCCCAGCTCTGTGGGGCCGT 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2523 AGTGGACAAGAGATAAGAGAGAACCGTGGGCTGGACAACACCTGGGCTTGGA 2582
```

```
Db   1321 AGTGGACAAGATAAGAGAGAACCGTGGCGTGGCGTGGACACCTTCAAAGTGACAGTGGGCGTGCA 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2583 CGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGTCCATCATGACAGCTGGCGTGGCGTGGACACCTTCAAAGTGACAGTGGGCGTGCA 2642

Db   1381 CGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGTCA 1440
          ||||||||||||||||||||||||||||||||||||||||
Qy   2643 TCTGGCTCCGAAATGTGACGTGTCCTTCCTGAAATCCGAGGACGGCAGTGGGAAGGGAGC 2702

Db   1441 TCTGGCTCCGAAATGTGACGTGTCCTTCCTGAATCCGAGGACGGCAGTGGGAAGGGAGC 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2703 AGCTCTCATCACTGCCGTGCCGTGGCCGTGCCGCCATCCGGGAGGACGGCAGTGGGGAAGGGAGC 2754

Db   1501 AGCTCTCATCACTGCCGTGCCGTGGCCGTGCCGCCATCCGGGAGGCTGGGCAGAGATAG 1552
```

RESULT 6

```
ACC098769
LOCUS       AC098769       231777 bp    DNA     linear   HTG 13-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-186P10, *** SEQUENCING IN PROGRESS
            ***, 16 unordered pieces.
ACCESSION   AC098769
VERSION     AC098769.10  GI:21735640
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 231777)
  AUTHORS   Muzny,D.M., Adams,C.J., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,J., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Puiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Soderegren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,D., Telrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,R.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 231777)
  AUTHORS   Worley,K.C.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (01-NOV-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 231777)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
  On Jul 12, 2002 this sequence version replaced gi:20467513.
COMMENT     Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            --------------- Project Information ---------------
            Center project name: GISF
            Center clone name: CH230-186P10
            --------------- Summary Statistics ---------------
            Sequencing vector: Plasmid;
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 216012 bases at least Q40
            Consensus quality: 217712 bases at least Q30
            Consensus quality: 218875 bases at least Q20
            --------------------------------------------------
            NOTE: Estimated insert size may differ from sequence length
            (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            NOTE: This is a 'working draft' sequence. It currently
            consists of 16 contigs. The true order of the pieces
            is not known and their order in this sequence record is
            arbitrary. Gaps between the contigs are represented as
            runs of N, but the exact sizes of the gaps are unknown.
            This record will be updated with the finished sequence
            as soon as it is available and the accession number will
            be preserved.
```

```
                  1      1530:  contig of 1530 bp in length
               1531      1630:  gap of unknown length
               1631      2737:  contig of 1107 bp in length
               2738      2837:  gap of unknown length
               2838      4352:  contig of 1515 bp in length
               4353      4452:  gap of unknown length
               4453      5558:  contig of 1106 bp in length
               5559      5658:  gap of unknown length
               5659      6941:  contig of 1283 bp in length
               6942      7041:  gap of unknown length
               7042      9713:  contig of 2672 bp in length
               9714      9813:  gap of unknown length
               9814     20278:  contig of 10465 bp in length
              20279     20378:  gap of unknown length
              20379     28422:  contig of 8044 bp in length
              28423     28522:  gap of unknown length
              28523     37797:  contig of 9275 bp in length
              37798     37897:  gap of unknown length
              37898     54558:  contig of 16661 bp in length
              54559     54658:  gap of unknown length
              54659     72827:  contig of 18169 bp in length
              72828     72927:  gap of unknown length
              72928    100350:  contig of 27423 bp in length
             100351    100450:  gap of unknown length
             100451    127127:  contig of 26677 bp in length
             127128    127227:  gap of unknown length
             127228    155687:  contig of 28460 bp in length
             155688    155787:  gap of unknown length
             155788    189137:  contig of 33350 bp in length
             189138    189237:  gap of unknown length
             189238    231777:  contig of 42540 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..231777
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="CH230-186P10"
```

```
RESULT 7
AC098769/c

Qy   1566 CACAG 1570
          |||||
Db  65651 CACAC 65655

Qy   1506 GACCCATGCGGTCGCGCCCTGTGAAGGATCGTCGTGTCCCACTCCAGATGG 1565
          || |||||||||||||||||| ||||||||||||||||||||||||||||
Db  65591 GACCGCATGCGGTCGCCCCCCTGTGAAGATGCTGCTGTGTGTGCCACTCCAGATGG 65650

Qy   1446 GCAGCTTCTGGAGGTTAAGAGAAGAAATGAAGTGGAAATGGAGCCAGAGGGTCTGAGCAAGGA 1505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  65531 GCAGCTTCTGGAGGTTAAGAGAAGAAATGAAGTGGAAATGGAGCCAGGGGTCTGAGCAAGGA 65590

Qy   1386 TCTGGCTCACCAACACACCGGGCCCCCAGAAGACCCTGGAGTCTCTGAAGCTGAGCCAGCACGA 1445
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  65471 TCTGGCTCACCAACACACCGGGCCCCCAGAAGACCCTGGAGTCTCTGAAGCTGAGCCAGCACGA 65530

Qy   1326 CCTCCGGCTCTGAGGATGGCACGCGACGGCCAAAGGGGGCTGCTATGGTGACCGGCGTTACCG 1385
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  65411 CCTCCGGCTCTGAGGATGGCACGCGACGGCCAAAGGGGGCTGCTATGGTGACCGGCGTTACCG 65470

Qy   1266 TTTTTGCCAAGCGTCTCCATAAGGCAGTGAGAGAGGCTGTGGCCGACTGTGANGTCCGGCTT 1325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  65351 TTTTTGCCAAGCGTCTCCATAAGGCAGTGAGAGAGGCTGTGGCCGACTGTGANGTCCGGCTT 65410
```

```
FEATURES
          *  54559   54658:  gap of unknown length
          *  54559   54827:  contig of 18169 bp in length
          *  72828   72927:  gap of unknown length
          *  72928  100350:  contig of 27423 bp in length
          * 100351  100450:  gap of unknown length
          * 100451  127127:  contig of 26677 bp in length
          * 127128  127227:  gap of unknown length
          * 127228  155687:  contig of 28460 bp in length
          * 155688  155787:  gap of unknown length
          * 155788  189137:  contig of 33350 bp in length
          * 189138  189237:  gap of unknown length
          * 189238  231777:  contig of 42540 bp in length
FEATURES          Location/Qualifiers
     source        1..231777
                   /organism="Rattus norvegicus"
                   /db_xref="taxon:10116"
                   /clone="CH230-186P10"

BASE COUNT   60544 a   52226 c   52675 g   62183 t    4149 others
ORIGIN

Query Match
Best Local Similarity  100.0%;   Score 187;  DB 2;  Length 231777;
Matches 187;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   2374  ACGGACTGCCTAGCCCTGCTACAGGTTCGTGCCATCCTGCCACCTAGGCTGTGGAGAGC  2433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 231437  AGCGACTGCCTAGCCCTGCTACAGGTTCGTGCCATCCTGCCACCTAGGCTGTGGAGAGC 231378

Qy   2434  ACGTGCCGATGACAGCATCATCGTGAAGAGAGGTGTGCACTGTGGTTGCCCCGGCCTGCA  2493
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 231377  ACGTGCCGATGACAGCATCATCGTGAAGAGAGGTGTGCACTGTGGTTGCCCCGGCCTGCA 231318

Qy   2494  CAGCTCTGTGCGCCAGCCATGCCGCCGCCGCGCCGTAGTCGACAAGATAAGAGAGACCGTGGCGCTG  2553
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 231317  CAGCTCTGTGCGCCAGCCATGCCGCCGCCGCGCCGTAGTGGACAAGATAAGAGAGAACCGTGGCGCTG 231258

Qy   2554  GACAACC  2560
           |||||||
Db 231257  GACAACC 231251
```

```
RESULT 8
LOCUS       MMU238540
MMU238540
LOCUS       MMU238540           5474 bp    mRNA      linear    ROD 04-FEB-2000
DEFINITION  Mus musculus mRNA for hexokinase II.
ACCESSION   AJ238540
VERSION     AJ238540.1  GI:6911968
KEYWORDS    hexokinase II; HKII gene.
SOURCE      Mus musculus
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 5474)
  AUTHORS   Heikkinen,S., Suppola,S., Malkki,M., Deeb,S.S., Janne,J. and
            Laakso,M.
  TITLE     Mouse hexokinase II gene: structure, cDNA, promoter analysis, and
            expression pattern
  JOURNAL   Mamm. Genome 11 (2), 91-96 (2000)
  MEDLINE   20122158
   PUBMED   10656921
REFERENCE   2  (bases 1 to 5474)
  AUTHORS   Heikkinen,S.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-1999) Heikkinen S.P., A.I.Virtanen Institute for
            Molecular Sciences, University of Kuopio, POBox 1627, 70211 Kuopio,
            FINLAND
COMMENT     Related sequences: Y11666, Y11667 & Y11668 (genomic DNA).
FEATURES          Location/Qualifiers
     source        1..5474
                   /organism="Mus musculus"
                   /db_xref="taxon:10090"
     source        1..1663
                   /organism="Mus musculus"
```

```
     source        1..5474
                   /strain="Balb/c x DBA/2"
                   /db_xref="taxon:10090"
                   /sex="male"
                   /tissue_type="hind limb muscle"
                   /dev_stage="5 week old"
     source        1664..5474
                   /organism="Mus musculus"
                   /strain="unknown"
                   /db_xref="taxon:10090"
                   /clone="HK2-26"
                   /tissue_type="striated muscle"
                   /dev_stage="adult"
                   /tissue_lib="mouse oligo-dT primed diaphragm cDNA library
                   from Stratagene Inc."
     gene          1..5453
                   /gene="HKII"
     5'UTR         <1..417
                   /gene="HKII"
     repeat_region 198..226
                   /gene="HKII"
                   /rpt_type=DIRECT
                   /note="A-rich region"
     CDS           418..3171
                   /gene="HKII"
                   /EC_number="2.7.1.1"
                   /function="glycolysis"
                   /codon_start=1
                   /product="hexokinase II"
                   /protein_id="CAB72257.1"
                   /db_xref="GI:6911969"
                   /translation="MIASHMIACLFTELNQNOVKDQVKDVQYLYHMRLSDETLLEISRRFR
                   KEMEKGLGATTHPTAAVKMLPTFVRSTPDGTEHGEFLALDLGGTNFRVLRVRVTDNGL
                   QRVEMENQIYAIPEDIMRGSGTOLFDHIAECLANFMDKLQIKEKKLPLGFTSFSPPCHQ
                   TKLDESFLVSWTKGFKSSGVEGRDVVDLIRKAIQRRGDFDIDIVAVVNDTVGTMMTCG
                   VDDQNCEIGLIVGTGSNACYMEEMRHIDMVEGDEGRMCINMEWGAFGDDGILNDIRTE
                   FDREIDMGSLNPGKQLFEKMISGMYMGELVRLILVKMAKAELLFQGKLSPELLTTGSF
                   ETKDVSDIEDDKDGIQKAYOILVRLGLSPLQEDCVATHRICQIVSTRSASLCAATLAA
                   VLWRIKENKGEERLRSTIGVDGSVYKKHPHFAKRLHKAVRRLVPDCDVRFLRSEDGSG
                   KGAAMVTAVAYRLADQHRAROKTLESLKLSHEQLLEVKRRMKVEMEQGLSKETHEAAP
                   VKMLPTVVCATPDGTEKGDFLALDLGGTNFRVLLVRVRNGKRGVEMHNKYISIPQEV
                   MHGTGEELFDHIVQCIADFLEYMGMKGVSLPLGFTFSFPCQONSLDQSILLKWTKGFK
                   ASGCEGEDVVTLLKEAIRRREEEFDLDVVAVVNDTVGTMMTCGYEDPHCEVGLIVGTGS
                   NACYMEEMRNVELVDGEEGRMCVNMEWGAFGDNGCLDDLRTVFDVAVDELSLNPGKQR
                   FEKMISGMYLGEIVRNILIDFTKRGLLFRGRISERLKTRGIFETKFLSQIESDCLALL
                   QVRAILRHLGLESTCDDSIIVKEVCTVVARRAAQLCGAGMAAVVDKIRENGLDNLKV
                   TVGVDGTLYKLHPHFAKVMHETVRDLAPKCDVSFLESEDGSGKGAALITAVACRIREA
                   GQR"
     variation     2358
                   /gene="HKII"
                   /note="silent variation compared to genomic sequence Acc#
                   Y11668 at location 21361"
                   /replace="t"
     3'UTR         3172..5453
                   /gene="HKII"
     repeat_region 3546..3578
                   /gene="HKII"
                   /note="GGA-repeat"
                   /rpt_type=DIRECT
     repeat_region 3803..3824
                   /gene="HKII"
                   /note="T-repeat"
                   /rpt_type=DIRECT
     polyA_signal  4706..4711
                   /gene="HKII"
     polyA_signal  4734..4739
                   /gene="HKII"
                   /evidence=not_experimental
                   /note="putative secondary signal"
     repeat_region 5157..5192
                   /gene="HKII"
                   /evidence=not_experimental
                   /note="putative secondary signal"
     variation     5157..5192
                   /gene="HKII"
                   /note="TG-repeat"
                   /rpt_type=DIRECT
                   /note="one extra GT repeat compared to genomic sequence
```

```
RESULT 9
LOCUS       MMHEXII3
MMHEXII3          28646 bp    DNA     linear   ROD 02-MAR-2000
DEFINITION  Mus musculus gene for hexokinase II, exon 3-18.
ACCESSION   Y11668
VERSION     Y11668.1  GI:1907085
KEYWORDS    hexokinase II.
SOURCE      Mus musculus.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 28646)
AUTHORS     Heikkinen,S., Suppola,S., Malkki,M., Deeb,S.S., Janne,J. and
            Laakso,M.
TITLE       Mouse hexokinase II gene: structure, cDNA, promoter analysis, and
            expression pattern
JOURNAL     Mamm. Genome 11 (2), 91-96 (2000)
MEDLINE     20122158
PUBMED      10656921
REFERENCE   2  (bases 1 to 28646)
AUTHORS     Heikkinen,S.
TITLE       Direct Submission
JOURNAL     Submitted (06-MAR-1997) S. Heikkinen, University of Kuopio,
            A.I.Virtanen Institute, PO Box 1627, SF-70211 Kuopio, FINLAND
FEATURES             Location/Qualifiers
     source          1..28646
                     /organism="Mus musculus"
                     /strain="129/SvJ"
                     /db_xref="taxon:10090"
                     /sex="female"
                     /tissue_type="liver"
                     /clone_lib="lambda FixII"
     intron          <1..2983
                     /number=2
     repeat_region   1992..2187
                     /partial
                     /rpt_family="SINE/B2"
                     /rpt_type=INVERTED
     repeat_region   2699..2830
```

```
Query Match    3.4%;  Score 95;  DB 10;  Length 5474;
Best Local Similarity   99.0%;  Pred. No. 6.5e-39;
Matches  195;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy   1246 GTCTACAGAGAAACATCCCATTTTTGCCAAAGCGTCTCCATAAGGCAGTGAGGAGGCTGGTG 1305
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1663 GTCTACAAGAGAAACATCCCCATTTTTGCCCAAGCGTCTCCATAAGGCAGTGAGGAGGCTGGTG 1722

Qy   1306 CCCGACTGTGATGTGTCCGCGTTCCTCCGCGTTCTGAGGATGGCAGCGGGCAAGGGGCTGCTATG 1365
          |||| |||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||
Db   1723 CCCGATTGTGATGTGTCCGCGTTCCTCCGCGTTCTGAGGATGGCAGCGCGCAAAGGGGCGCTATG 1782

Qy   1366 GTGACGGCGCGGTGGCGCTTACCGTCTCGGCTGACCAACACCACCGGGGCAGAAGAGACCCTGGAG 1425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1783 GTGACGGCGCGGTGGCGCTTACCGTCTCGGCTGACCAACACCACCGGGGCCCCGCCAGAGAAGAGACCCTGGAG 1842

Qy   1426 TCTCTGAACGCTGACCCA 1442
          ||||||||||||||||||
Db   1843 TCTCTGAACGCTGAGGCCA 1859
```

```
ORIGIN
BASE COUNT   1320 a   1268 c   1561 g   1325 t

     misc_feature    5454..5474
                     /gene="HKII"
                     /note="partial polyA tail"
     polyA_site      5453
                     /gene="HKII"
                     /evidence=experimental
     polyA_signal    5431..5436
                     /gene="HKII"
                     /note="primary signal"
                     /evidence=experimental
```

Acc# Y11668 at locations 27179..27212"

```
     exon                       /rpt_type=DIRECT
                     /rpt_unit=2699..2702
                     2984..3132
                     /product="hexokinase II"
                     /number=3
                     /usedin=Y11666:hexIImrna
                     /usedin=Y11666:hexIICDS
     intron          3133..7348
                     /number=3
     repeat_region   4446..4497
                     /rpt_type=DIRECT
                     /rpt_unit=4446..4447
     repeat_region   6402..6441
                     /rpt_type=DIRECT
                     /rpt_unit=6402..6403
     exon            7349..7468
                     /product="hexokinase II"
                     /number=4
                     /usedin=Y11666:hexIImrna
                     /usedin=Y11666:hexIICDS
     intron          7469..8210
                     /number=4
     repeat_region   7920..7959
                     /rpt_type=DIRECT
                     /rpt_unit=7920..7921
     exon            8211..8306
                     /product="hexokinase II"
                     /number=5
                     /usedin=Y11666:hexIImrna
                     /usedin=Y11666:hexIICDS
     intron          8307..8882
                     /number=5
     exon            8883..8982
                     /product="hexokinase II"
                     /number=6
                     /usedin=Y11666:hexIImrna
                     /usedin=Y11666:hexIICDS
     intron          8983..9233
                     /number=6
     repeat_region   9234..9417
                     /partial
                     /note="B1-F-repeat"
                     /rpt_family="SINE/Alu"
                     /rpt_type=INVERTED
     exon            9418..12798
                     /number=7
     intron          9668..9789
                     /partial
                     /note="B1-MM-repeat"
                     /rpt_family="SINE/Alu"
                     /rpt_type=INVERTED
     repeat_region   11156..11299
                     /partial
                     /note="B1-MM-repeat"
                     /rpt_family="SINE/Alu"
                     /rpt_type=INVERTED
     repeat_region   11894..12094
                     /partial
                     /note="L1-MM-repeat"
                     /rpt_family="LINE/L1"
     repeat_region   12218..12363
                     /partial
                     /rpt_family="LINE/L1"
                     /rpt_type=DIRECT
     exon            12799..12954
                     /product="hexokinase II"
                     /number=8
                     /usedin=Y11666:hexIImrna
                     /usedin=Y11666:hexIICDS
     intron          12955..13979
                     /number=8
```

```
repeat_region      13046. .13083
                   /rpt_type=INVERTED
repeat_region      13873. .13892
                   /rpt_unit=13046. .13047
                   /rpt_type=INVERTED
exon               13980. .14213
                   /rpt_unit=13873. .13874
                   /product="hexokinase II"
                   /number=9
                   /usedin=Y11666:hexIImrna
intron             14214. .15661
                   /usedin=Y11666:hexIICDS
exon               15662. .15966
                   /number=9
                   /usedin=Y11666:hexIImrna
                   /product="hexokinase II"
                   /number=10
intron             15967. .17106
                   /usedin=Y11666:hexIImrna
                   /usedin=Y11666:hexIICDS
repeat_region      15988. .16050
                   /number=10
                   /rpt_type=INVERTED
exon               17107. .17255
                   /rpt_unit=15988. .15989
                   /product="hexokinase II"
                   /number=11
                   /usedin=Y11666:hexIImrna
intron             17256. .17534
                   /usedin=Y11666:hexIICDS
exon               17535. .17654
                   /number=11
                   /usedin=Y11666:hexIImrna
                   /product="hexokinase II"
                   /number=12
repeat_region      17655. .20306
                   /usedin=Y11666:hexIImrna
                   /usedin=Y11666:hexIICDS
                   /number=12
intron             17757. .17897
                   /partial
repeat_region      19253. .19896
                   /note="B1-MM-repeat"
                   /rpt_family="SINE/Alu"
                   /rpt_type=DIRECT
                   /partial
repeat_region
                   /note="L1-MM-repeat"
                   /rpt_family="LINE/L1"
                   /rpt_type=INVERTED
                   /product="hexokinase II"
                   /number=13
                   /partial
exon               20307. .20402
                   /usedin=Y11666:hexIImrna
                   /rpt_family="LINE/L1"
                   /rpt_type=INVERTED
                   /number=13
intron             20762. .20930
                   /usedin=Y11666:hexIImrna
                   /rpt_unit=20608. .20609
                   /partial
repeat_region      20608. .20642
                   /note="B2-repeat"
                   /rpt_family="SINE/B2"
                   /rpt_type=INVERTED
                   /number=13
exon               21356. .21455
                   /product="hexokinase II"
                   /number=14
                   /usedin=Y11666:hexIImrna
intron             21456. .21551
                   /usedin=Y11666:hexIICDS
exon               21552. .21735
                   /product="hexokinase II"
                   /number=15
                   /usedin=Y11666:hexIImrna
```

```
intron             21736. .22802
                   /usedin=Y11666:hexIICDS
repeat_region      21949. .22005
                   /number=15
                   /rpt_type=DIRECT
repeat_region      22603. .22630
                   /rpt_type=INVERTED
exon               22803. .22958
                   /rpt_unit=22603. .22604
                   /product="hexokinase II"
                   /number=16
                   /usedin=Y11666:hexIImrna
intron             22959. .23557
                   /usedin=Y11666:hexIICDS
exon               23558. .23791
                   /number=16
```

```
Query Match
Best Local Similarity 100.0%;  Pred. No. 1.4e-32;
Matches  83; Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
Score 83;  DB 10;  Length 28646;

Qy   1360 GCTATGGTGACGGCGGTGGCGTTACCGTCTGGCTGACCAACACACCGGGGCCCCGCCAGGAAGACC 1419
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  15756 GCTATGGTGACGGCGGTGGCGTTACCGTCTGGCTGACCAACACACCGGGGCCCCGCCAGGAAGACC 15815

Qy   1420 CTGGAGTCTCTGAAGCTGAGCCA 1442
          |||||||||||||||||||||||
Db  15816 CTGGAGTCTCTGAAGCTGAGCCA 15838

RESULT 10
AC116811/c
LOCUS       AC116811              67493 bp    DNA     linear    HTG 20-MAY-2002
DEFINITION  Mus musculus clone RP24-356D15, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC116811
VERSION     AC116811.2  GI:20986619
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Mus musculus
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 67493)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP24-356D15
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 67493)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Dewar,K., Diaz,J.S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Illev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talanas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Triglijo,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 67493)
```

AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
          Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
          Boukhgalter,B., Brown,A., Camarata,J., Camppiano,A., Chang,J.,
          Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
          Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
          Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
          Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
          Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
          Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
          Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K.,
          Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
          Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
          Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
          Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
          O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
          Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
          Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
          Schupbach,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
          Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
          Theodore,J., Topham,K., Travers,M., Travis,N., Triglilo,J.,
          Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
          Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE   Direct Submission
JOURNAL   Submitted (20-MAY-2002) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT   On May 20, 2002 this sequence version replaced gi:19882065.
          All repeats were identified using RepeatMasker:
          Smit,A.F.A. & Green,P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          --------------Genome Center
          Center: Whitehead Institute/ MIT Center for Genome Research
          Center code: WIBR
          Web site: http://www-seq.wi.mit.edu
          Contact: sequence_submissions@genome.wi.mit.edu
          --------------Project Information
          Center project name: L25559
          Center clone name: 356_D.15

          NOTE: This record contains 83 individual
          sequencing reads that have not been assembled into
          contigs. Runs of N are used to separate the reads
          and the order in which they appear is completely
          arbitrary. Low-pass sequence sampling is useful for
          identifying clones that may be gene-rich and allows
          overlap relationships among clones to be deduced.
          However, it should not be assumed that this clone
          will be sequenced to completion. In the event that
          the record is updated, the accession number will
          be preserved.

          *   1 707: contig of 707 bp in length
          *   708 807: gap of 100 bp
          *   808 1524: contig of 717 bp in length
          *   1525 1624: gap of 100 bp
          *   1625 2360: contig of 736 bp in length
          *   2361 2460: gap of 100 bp
          *   2461 3170: contig of 710 bp in length
          *   3171 3270: gap of 100 bp
          *   3271 3990: contig of 720 bp in length
          *   3991 4090: gap of 100 bp
          *   4091 4804: contig of 714 bp in length
          *   4805 4904: gap of 100 bp
          *   4905 5580: contig of 676 bp in length
          *   5581 5680: gap of 100 bp
          *   5681 6413: contig of 733 bp in length
          *   6414 6513: gap of 100 bp
          *   6514 7211: contig of 698 bp in length
          *   7212 7311: gap of 100 bp
          *   7312 8039: contig of 728 bp in length
          *   8040 8139: gap of 100 bp
          *   8140 8852: contig of 713 bp in length
          *   8853 8952: gap of 100 bp
          *   8953 9685: contig of 733 bp in length
          *   9686 9785: gap of 100 bp

          *   9786 10489: contig of 704 bp in length
          *   10490 10589: gap of 100 bp
          *   10590 11296: contig of 707 bp in length
          *   11297 11396: gap of 100 bp
          *   11397 12119: contig of 723 bp in length
          *   12120 12219: gap of 100 bp
          *   12220 12933: contig of 714 bp in length
          *   12934 13033: gap of 100 bp
          *   13034 13750: contig of 717 bp in length
          *   13751 13850: gap of 100 bp
          *   13851 14568: contig of 718 bp in length
          *   14569 14668: gap of 100 bp
          *   14669 15373: contig of 705 bp in length
          *   15374 15473: gap of 100 bp
          *   15474 16198: contig of 725 bp in length
          *   16199 16298: gap of 100 bp
          *   16299 17011: contig of 713 bp in length
          *   17012 17111: gap of 100 bp
          *   17112 17830: contig of 719 bp in length
          *   17831 17930: gap of 100 bp
          *   17931 18646: contig of 716 bp in length
          *   18647 18746: gap of 100 bp
          *   18747 19479: contig of 733 bp in length
          *   19480 19579: gap of 100 bp
          *   19580 20297: contig of 718 bp in length
          *   20298 20397: gap of 100 bp
          *   20398 21106: contig of 709 bp in length
          *   21107 21206: gap of 100 bp
          *   21207 21931: contig of 725 bp in length
          *   21932 22031: gap of 100 bp
          *   22032 22743: contig of 712 bp in length
          *   22744 22843: gap of 100 bp
          *   22844 23575: contig of 732 bp in length
          *   23576 23675: gap of 100 bp
          *   23676 24392: contig of 717 bp in length
          *   24393 24492: gap of 100 bp
          *   24493 25213: contig of 721 bp in length
          *   25214 25313: gap of 100 bp
          *   25314 26035: contig of 722 bp in length
          *   26036 26135: gap of 100 bp
          *   26136 26857: contig of 722 bp in length
          *   26858 26957: gap of 100 bp
          *   26958 27669: contig of 712 bp in length
          *   27670 27769: gap of 100 bp
          *   27770 28498: contig of 729 bp in length
          *   28499 28598: gap of 100 bp
          *   28599 29306: contig of 708 bp in length
          *   29307 29406: gap of 100 bp
          *   29407 30111: contig of 705 bp in length
          *   30112 30211: gap of 100 bp
          *   30212 30917: contig of 706 bp in length
          *   30918 31017: gap of 100 bp
          *   31018 31716: contig of 699 bp in length
          *   31717 31816: gap of 100 bp
          *   31817 32541: contig of 725 bp in length
          *   32542 32641: gap of 100 bp
          *   32642 33357: contig of 716 bp in length
          *   33358 33457: gap of 100 bp
          *   33458 34167: contig of 710 bp in length
          *   34168 34267: gap of 100 bp
          *   34268 34991: contig of 724 bp in length
          *   34992 35091: gap of 100 bp
          *   35092 35819: contig of 728 bp in length
          *   35820 35919: gap of 100 bp
          *   35920 36642: contig of 723 bp in length
          *   36643 36742: gap of 100 bp
          *   36743 37445: contig of 703 bp in length
          *   37446 37545: gap of 100 bp
          *   37546 38270: contig of 725 bp in length
          *   38271 38370: gap of 100 bp
          *   38371 39096: contig of 726 bp in length
          *   39097 39196: gap of 100 bp
          *   39197 39905: contig of 709 bp in length

```
 *    39906 40005: gap of         100 bp in length
 *    40006 40716: contig of 711 bp in length
 *    40717 40816: gap of         100 bp
 *    40817 41551: contig of 735 bp in length
 *    41552 41651: gap of         100 bp
 *    41652 42366: contig of 715 bp in length
 *    42367 42466: gap of         100 bp
 *    42467 43181: contig of 714 bp in length
 *    43181 43280: gap of         100 bp
 *    43281 44004: contig of 724 bp in length
 *    44005 44104: gap of         100 bp
 *    44104 44832: contig of 728 bp in length
 *    44105 44833 44932: gap of   100 bp
 *          44933 45649: contig of 717 bp in length
 *    45649: gap of   100 bp in length

Query Match          3.0%;   Score 83;   DB 2;   Length 67493;
Best Local Similarity   100.0%;   Pred. No. 1.2e-32;
Matches  83;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1360 GCTATGGTGACGGCGGGTGGCTTACCGTCTGGCTGACCAACACCGGGCGCCCGCCCAGAAGACC 1419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   31667 GCTATGGTGACGGCGGGTGGCTTACCGTCTGGCTGACCAACACCGGGCGCCCGCCCAGAAGACC 31608

Qy   1420 CTGGAGTCTCTGAAGCTGAGCCA 1442
          |||||||||||||||||||||||
Db   31607 CTGGAGTCTCTGAAGCTGAGCCA 31585
```

```
RESULT 11
AC007305/c
LOCUS         AC007305           214839 bp    DNA    linear    HTG 15-JUN-2002
DEFINITION    Mus musculus chromosome 6 clone mgs1-211120 strain 129/SvJ, WORKING
              DRAFT SEQUENCE, 2 ordered pieces.
ACCESSION     AC007305
VERSION       AC007305.33  GI:21427706
KEYWORDS      HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE        Mus musculus (house mouse)
  ORGANISM    Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1  (bases 1 to 214839)
  AUTHORS     Chen,F., Tilahun,Y., Meisler,M. and Roe,B.A.
  TITLE       Mus musculus Chromosome 6 BAC Clone mgs1-211120 In MND2 Region
  JOURNAL     Unpublished
REFERENCE     2  (bases 1 to 214839)
  AUTHORS     Chen,F., Tilahun,Y., Meisler,M. and Roe,B.A.
  TITLE       Direct Submission
  JOURNAL     Submitted (14-APR-1999) Department of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
REFERENCE     3  (bases 1 to 214839)
  AUTHORS     Chen,F., Tilahun,Y., Meisler,M. and Roe,B.A.
  TITLE       Direct Submission
  JOURNAL     Submitted (15-JUN-2002) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
              On Jun 15, 2002 this sequence version replaced gi:19387636.
COMMENT       -------------------- Genome Center
              Center: Department Of Chemistry And Biochemistry
              The University of Oklahoma
              Center code:UOKNOR
              -------------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submitter.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            * The finished sequence will contain 9488 bp in length
            *        1 9488: contig of 9488 bp in length
            *     9489 94589 94588: gap of unknown length
```

```
FEATURES             Location/Qualifiers
     *         94589 214839: contig of 120251 bp in length.
     source          1..214839
                     /organism="Mus musculus"
                     /strain="129/SvJ"
                     /db_xref="taxon:10090"
                     /chromosome="6"
                     /clone="mgs1-211120"
                     /clone_lib="Genome Systems mouse ES cell BAC libraries I
                     and II"
BASE COUNT    59644 a   50880 c   48896 g   55319 t    100 others
ORIGIN
```

```
Query Match          3.0%;   Score 83;   DB 2;   Length 214839;
Best Local Similarity   100.0%;   Pred. No. 9.8e-33;
Matches  83;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1360 GCTATGGTGACGGCGGGTGGCTTACCGTCTGGCTGACCAACACCGGGCGCCCGCCCAGAAGACC 1419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2172 GCTATGGTGACGGCGGGTGGCTTACCGTCTGGCTGACCAACACCGGGCGCCCGCCCAGAAGACC 2113

Qy   1420 CTGGAGTCTCTGAAGCTGAGCCA 1442
          |||||||||||||||||||||||
Db   2112 CTGGAGTCTCTGAAGCTGAGCCA 2090
```

```
RESULT 12
MMHEXII2
LOCUS         MMHEXII2             404 bp    DNA    linear    ROD 02-MAR-2000
DEFINITION    Mus musculus gene for hexokinase II, exon 2.
ACCESSION     Y11667
VERSION       Y11667.1  GI:1907084
KEYWORDS      hexokinase II.
SOURCE        Mus musculus.
  ORGANISM    Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1  (bases 1 to 404)
  AUTHORS     Heikkinen,S., Suppola,S., Malkki,M., Deeb,S.S., Janne,J. and
              Laakso,M.
  TITLE       Mouse hexokinase II gene: structure, cDNA, promoter analysis, and
              expression pattern
  JOURNAL     Mamm. Genome 11 (2), 91-96 (2000)
  MEDLINE     20122158
  PUBMED      10656921
REFERENCE     2  (bases 1 to 404)
  AUTHORS     Heikkinen,S.
  TITLE       Direct Submission
  JOURNAL     Submitted (06-MAR-1997) S. Heikkinen, University of Kuopio,
              A.I.Virtanen Institute, PO Box 1627, SF-70211 Kuopio, FINLAND
FEATURES             Location/Qualifiers
     source          1..404
                     /organism="Mus musculus"
                     /strain="129/SvJM"
                     /db_xref="taxon:10090"
                     /sex="female"
                     /tissue_type="liver"
                     /clone_lib="lambda FixII"
     intron          <1..156
                     /number=1
     exon            157..319
                     /number=2
                     /product="hexokinase II"
                     /usedin=Y11666:hexIImrna
                     /usedin=Y11666:hexIICDS
     intron          320..>404
                     /number=2
     repeat_region   324..355
                     /rpt_type=INVERTED
                     /rpt_unit=324..325
BASE COUNT       91 a    98 c    105 g    110 t
ORIGIN
```

RESULT 13

```
Query Match    Best Local Similarity  100.0%;  Score 72;  DB 10;  Length 404;
Matches   72;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   75  TCTCTACCACATGCGTCTCTCAGATGAGAGACCCTTCTGGAGATTTGTAGCGGTGCGGGTTCCGGAA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168  TCTCTACCACATGCGTCTCTCAGATGAGAGACCCTTCTGGAGATTTCTAGGCGGTGCGGGTTCCGGAA 227

Qy  135  GGGAGATGCAGAA 146
         ||||||||||||
Db  228  GGGAGATGCAGAA 239
```

LOCUS   RATHK2

LOCUS   RATHK2   1731 bp   DNA   linear   ROD 07-NOV-2001
DEFINITION   Rattus norvegicus HK2 gene for type II hexokinase, exon1 and
             promoter region.
ACCESSION   D26393
VERSION   D26393.1 GI:893403
KEYWORDS   .
SOURCE   Rattus norvegicus
  ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE   1 (bases 1 to 1731)
  AUTHORS   Ichihara,J., Shinohara,Y., Kogure,K. and Terada,H.
  TITLE   Nucleotide sequence of the 5'-flanking region of the rat type II
             hexokinase gene
  JOURNAL   Biochim. Biophys. Acta 1260 (3), 365-368 (1995)
   MEDLINE   95178553
REFERENCE   2 (bases 1 to 1731)
  AUTHORS   Shinohara,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (04-JAN-1994) Yasuo Shinohara, University of Tokushima,
             Faculty of Pharmaceutical Sciences; 1 Shomachi, Tokushima,
             Tokushima 770, Japan (E-mail:yasuo@ph.tokushima-u.ac.jp,
             Tel:81-886-33-7278, Fax:81-886-33-5196)
FEATURES   Location/Qualifiers
     source   1..1731
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
     misc_signal   56..61
                     /note="CF1 (transcription factor binding site)"
     misc_signal   185..190
                     /note="CF1 (transcription factor binding site)"
     misc_signal   464..469
                     /note="PEA3 (transcription factor binding site)"
     misc_signal   497..503
                     /note="PuF (transcription factor binding site)"
     misc_signal   844..856
                     /note="AP2.PuF.AP2 (transcription factor binding site)"
     gene   938..943
     GC_signal   1073..1077
     CAAT_signal   1101..1106
     GC_signal   1129..1133
     TATA_signal   1159..1683
     exon   number=1
     gene   /evidence=experimental
                     1621..1683
                     /gene="HK2"
     CDS   1621..>1683
                     /gene="HK2"
                     /codon_start=1
                     /EC_number="2.7.1.1"
                     /product="type II hexokinase"
                     /protein_id="BAA05409.1"
                     /db_xref="GI:1525568"
                     /translation="MIASHMIACLFTELNQQVQK"
BASE COUNT   385 a   491 c   474 g   381 t
               1696..1701
```

RESULT 14

```
Query Match    Best Local Similarity  100.0%;  Score 65;  DB 10;  Length 1731;
Matches   65;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1  ATGATCGCCTCGCCATATGAATGGCCTCGTCGTTATTCACGGAGCGTCAACCAAAAACCAAGTGCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621  ATGATCGCCTCGCCATATGAATGGCCTCGTCGTTATTCACGGAGCGTCAACCAAAAACCAAGTGCAG 1680

Qy   61  AAGGT 65
         |||||
Db 1681  AAGGT 1685
```

LOCUS   AY082375

LOCUS   AY082375   5067 bp   DNA   linear   ROD 26-MAR-2002
DEFINITION   Rattus norvegicus strain Sprague-Dawley type II hexokinase gene,
             exon 1 and partial cds.
ACCESSION   AY082375
VERSION   AY082375.1 GI:19743673
KEYWORDS   .
SOURCE   Rattus norvegicus
  ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE   1 (bases 1 to 5067)
  AUTHORS   Rempel,A.
  TITLE   Normal type II hexokinase promoter, first exon, and first intron
             from hepatocytes
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 5067)
  AUTHORS   Rempel,A.
  TITLE   Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Biological Chemistry, Johns Hopkins
             University, 725 N. Wolfe Street, Baltimore, MD 21205, USA
FEATURES   Location/Qualifiers
     source   1..5067
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /cell_type="hepatocyte"
     mRNA   4288..>4810
                     /product="type II hexokinase"
     exon   4288..4810
                     /number=1
     CDS   4748..>4810
                     /codon_start=1
                     /product="type II hexokinase"
                     /protein_id="AAL92551.1"
                     /db_xref="GI:19743674"
                     /translation="MIASHMIACLFTELNQQVQK"
BASE COUNT   1228 a   1244 c   1254 g   1341 t
ORIGIN
```

RESULT 15

```
Query Match    Best Local Similarity  100.0%;  Score 65;  DB 10;  Length 5067;
Matches   65;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1  ATGATCGCCTCGCCATATGAATCGCCTCGTCGTTATTCACGGAGCTCAACCAAAAACCAAGTGCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4748  ATGATCGCCTCGGCATATGAATCGCCTCGTCGTTATTCACGGAGCTCAACCAAAAACCAAGTGCAG 4807

Qy   61  AAGGT 65
         |||||
Db 4808  AAGGT 4812
```

LOCUS   RNU19605

LOCUS   RNU19605   5150 bp   DNA   linear   ROD 03-MAY-2001

```
DEFINITION   Rattus norvegicus type II hexokinase gene, partial cds and promoter
             region.
ACCESSION    U19605
VERSION      U19605.2  GI:13937421
KEYWORDS     .
SOURCE       Rattus norvegicus.
  ORGANISM   Rattus norvegicus.
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE    1  (bases 1 to 5150)
  AUTHORS    Mathupala,S.P., Rempel,A. and Pedersen,P.L.
  TITLE      Glucose catabolism in cancer cells. Isolation, sequence, and
             activity of the promoter for type II hexokinase
  JOURNAL    J. Biol. Chem. 270 (28), 16918-16925 (1995)
  MEDLINE    95348123
   PUBMED    7622509
REFERENCE    2  (bases 1 to 5150)
  AUTHORS    Mathupala,S.P.
  TITLE      Direct Submission
  JOURNAL    Submitted (09-JAN-1995) Biological Chemistry, Johns Hopkins
             University School of Medicine, 725 N. Wolfe Street, Baltimore, MD
             21205, USA
REFERENCE    3  (bases 1 to 5150)
  AUTHORS    Lee,M.G., Ko,Y.H. and Pedersen,P.L.
  TITLE      Direct Submission
  JOURNAL    Submitted (03-MAY-2001) Biological Chemistry, Johns Hopkins
             University School of Medicine, 725 N. Wolfe Street, Baltimore, MD
             21205, USA
  REMARK     Sequence update by submitter
COMMENT      On May 3, 2001 this sequence version replaced gi:901870.
FEATURES             Location/Qualifiers
     source          1..5150
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="29-1/XbaI"
                     /sex="female"
                     /cell_line="AS-30D"
                     /tissue_type="hepatoma, ascites"
                     /clone_lib="Lambda-FixII/AS-30D genomic"
     CDS             4831..>4893
                     /codon_start=1
                     /product="type II hexokinase"
                     /protein_id="AAB09025.1"
                     /db_xref="GI:1568599"
                     /translation="MIASHMIACLFTELNQNQVQK"
     exon            4340..4893
                     /number=1
     TATA_signal     4284..4288
     CAAT_signal     4340..4344
     intron          4894..>5150
                     /number=1
BASE COUNT     1249 a   1263 c   1277 g   1361 t
ORIGIN
```

```
Query Match    2.4%;  Score 65;  DB 10;  Length 5150;
Best Local Similarity  100.0%;  Pred. No. 7.9e-23;
Matches  65;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1 ATGATCGGCTCGCATATGATCGGCTCGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4831 ATGATCGGCTCGCATATGATCGGCTCGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG 4890

Qy      61 AAGGT 65
           |||||
Db    4891 AAGGT 4895
```

Search completed: May 27, 2003, 16:34:47
Job time : 8826 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: May 27, 2003, 11:57:02 ; Search time 596 Seconds
(without alignments)
10406.046 Million cell updates/sec

Title: US-09-808-743A-1
Perfect score: 2754
Sequence: 1 atgatcgcctcgcatatgat.........gggaggctgggcaagagtag 2754

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size :   50

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :   N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2754	100.0	2754	24	AAI67594	Rat mutant type II
2	2753	100.0	2770	18	AAT78599	AS-30D tumour Type
3	2550	92.6	3635	18	AAV00125	Rat hexokinase II
4	2550	92.6	3635	18	AAV00086	Rat hexokinase II
5	102	3.7	299	18	AAT80522	AS-30D tumour Type
6	65	2.4	5150	18	AAT78598	AS-30D tumour Type

ALIGNMENTS

RESULT 1
AAI67594

ID   AAI67594 standard; DNA; 2754 BP.
XX
AC   AAI67594;
XX
DT   27-FEB-2002 (first entry)
XX
DE   Rat mutant type II hexokinase encoding DNA.
XX
KW   Hexokinase; cell proliferation; glycolytic tumour; cancer; mutant;
KW   type II hexokinase; cytostatic; glycolysis inhibitor; gene therapy; ds.
XX
OS   Rattus norvegicus.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..2754
FT   /*tag= a
FT   /product= "mutant type II hexokinase"
XX
PN   WO200168667-A1.
XX
PD   20-SEP-2001.
XX
PF   14-MAR-2001; 2001WO-US08335.
XX
PR   14-MAR-2000; 2000US-189222P.
XX
PA   (UJJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI   Pedersen PL, Mathupala SP;
XX
DR   WPI; 2002-049005/06.
DR   P-PSDB; AAG660000.
XX
PT   Inhibiting proliferation of highly glycolytic tumors, e.g. in gastric
PT   cancer, hepatoma, colorectal cancer or lung cancer, by contacting cells
PT   with antisense molecules that hybridize with a nucleic acid encoding
PT   hexokinase -
XX
PS   Claim 2; Fig 6; 59pp; English.
XX
CC   The invention provides a method for inhibiting proliferation of tumour
CC   cells characterized by having a highly glycolytic phenotype. The method
CC   involves contacting the cells with an antisense polynucleotide or
CC   oligonucleotide that hybridizes with a mRNA encoding a hexokinase under
CC   conditions that allow hybridization of the antisense polynucleotide with
CC   the mRNA, thus inhibiting the proliferation of tumour cells. The method
CC   is useful for inhibiting the expression of a hexokinase in highly glycolytic
CC   for modulating the cellular proliferative disorder comprises low
CC   tumours. In particular, the cellular proliferative disorder comprises low
CC   grade astrocytoma, anaplastic astrocytoma, glioblastoma, medulloblastoma,
CC   gastric cancer, hepatoma, colorectal cancer, colorectal adenoma, acute
CC   myelogenous leukemia, lung cancer, renal cancer, leukemia, breast cancer,
CC   prostate cancer, endometrial cancer, bone cancer, squamous cell cancer
CC   and neuroblastoma. The present sequence represents the DNA encoding a rat
CC   mutant type II hexokinase.
XX
SQ   Sequence 2754 BP; 642 A; 659 C; 869 G; 584 T; 0 other;

Query Match   100.0%;   Score 2754;   DB 24;   Length 2754;
Best Local Similarity   100.0%;
Matches 2754;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG 60
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG 60

Qy   61 AAGGTTGACCAATTTCTCTCTACCACATGCCTCTTCTCAGATGAGAGACCCTTCTGGAGATTTCT 120
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAGGTTGACCAATTTCTCTCTACCACATGCCTCTTCTCAGATGAGAGACCCTTCTGGAGATTTCT 120

```
Qy    121 AGGCGGTTCCGGAAGGAGGATGAGAAAGGGGCTAGGAGCTACCAGCCACCAGCAGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGGCGGTTCCGGAAGGAGGATGAGAAAGGGGCTAGGAGCTACCAGCCACCAGCAGCT 180

Qy    181 GTGAAAATGTTGCCTACCTTTGTGAGGTCAACATGGGGGAGTTC 240
          |||||||||||||||||||||||||||||||||
Db    181 GTGAAAATGTTGCCTACCTTTGTGAGGTCAACATGGGGGAGTTC 240

Qy    241 CTGGCTCTTGGATCTTGGAGGAAACCAAACTTCCGTTGTGACGGACAAT 300
          ||||||||||||||||||||
Db    241 CTGGCTCTTGGATCTTGGAGGAAACCAAACTTCCGTTGTGACGGACAAT 300

Qy    301 GGCCTTCCACGAGAGTGGGAGATGGAGAGAACCAACTTGGAGGTAAGGTGACGGACAAT 360
          |||||||||||
Db    301 GGCCAGTGGAACCCAGCTGTTTGACCACACTGGCCAACTTCATGACAAG 360

Qy    361 GGCAGTGGAACCCAGCTGTTTGACCACACTGGCCAACTTCATGACAAG 420
          ||||||
Db    361 ...

Qy    421 CTACAAATCAAAGAGAGAGCTCCCTTCTCGGTTTCCCCTGCCACCAG 480

Qy    481 ACAAAAACTGGATGAGAGTTTTTTGGTTCTCGGTTCAAGTCCAGTGCGTG 540

Qy    541 GAAGGCAGAGAGTGTGGTGACCTGAGGTTTATTCCAGCGCGAGGGACTTTGAC 600

Qy    601 ATTGACATTGTGCCGTGTGGGTTGGGACCATGAATGACTTGTGGGCTATGAT 660

Qy    661 GATCAGAAACTCCGAGATTGGTCTCATTGTGTGGGCCACTGGCACCAACCCTGCCACCAAGGGAG 720

Qy    721 GAAATGCGTCATATATTGACACATGGTGGAGGGGCCCATGTCCATCAAACATGGAG 780

Qy    781 TGGGGAGCCTTTGGGGACGACGGGTACACTCAATGACATCCGAACCAAGATGTTTGACCGAGAG 840

Qy    841 ATCGACATGGGCTCGCTGAACCCTGGAAGCAGCTGTTGTTTGAGAGAGTGATTAGCGGGATG 900

Qy    901 TACATGGGGGAGCTGGTCTCAGCCTCATCCTCCTGGTGAAGATGGCCAAGGCAGCTGTTGTTC 960

Qy    961 CAAGGGAAAACTCAGCCCAGAAACTCCTTACCACTGGCTCTCCTTCGAGAGCAAAGATGTCTCG 1020

Qy   1021 GATATTGAAGAGGAGGGGCTACCAAAATCCTGATCGCGCCTGGGT 1080

Qy   1081 CTGAATCCATTGCAGGAGGATTGTGTGGCCACGCAACTGCAGATCTGTGTGTCCACG 1140

Qy   1141 CGCTCGGCCCAGTCTGTTGGCGACCAGCCCACCCTGCGCGGTGCTGCTGTGGGCAAATCAAAGAGAAAC 1200
```

(column 2)

```
Qy   1201 AAGGGCGAGGAGCGACTTCCGCTCCACCACCATCCGTGTCGTGATGGTGGCTCCGTCTCACAAGAAACAT 1260

Qy   1261 CCCCATTTTGCCAAGGCGTCTCCATAAGGCACTGAGGAGGTCGGTGCCGACTGTGATGTC 1320

Qy   1321 CGCTTCCTCGCGGATGGGAGCCAACGGGTGCTGCTATGGTGGTGCCCGACGGGGGCT 1380

Qy   1381 TACCGTCTGGGCTGGTGACCAAACACCGGGCTGCTGTTATGGTGATCGGCTGAAGCTGAGC 1440

Qy   1441 CACGAGCAGCCTTGGGGTGGACAAACACCGGCCCCCGGCCCAGAAGAACGGGGGCTTGCTT 1500

Qy   1501 AAGGAGAACGCATGCGTCGCGCCCGTGTGAAAGAAAGATGACACACTTACGTGTCACTCCA 1560

Qy   1561 GATGGGCACAGAAGAAGACTTCTTGGCCCTTGGAGAATGAAGGTCGGAAATGGCAGACGGG 1620

Qy   1621 CTGCTGGTGGTGCGGTGTAATGGCATGCACTGGGCGAACGGGAAAGGAAGCTCTTTGGGTC 1680

Qy   1681 TCCATCCCACAGGAGGTTATGCATGGCGAACTGGTCCCTGCCTTTGGTGTGCCTTCGGGTC 1740

Qy   1741 TGCATTCGGGACTTCCTGGAGTTACATGGCATGCCTGGTGTGACCGAGATCTGA 1800

Qy   1801 ACATTCTCGTTCCCTTGCCAGAGAACAGCCTAGACCAGAGACATCCTCAAGTGGACA 1860

Qy   1861 AAGGGATTCAAGGCATCTGGCTGCCGGAGGGTGAGGATGGTGGCTCACCTTGCTGAGGAAGCG 1920

Qy   1921 ATTCACCGCGGCAGAGAGTTTGACCTGGATGTGTTGGCCGTGGTGAATGACACAGTTGGG 1980

Qy   1981 ACTATGATGACTTGTGGGCTACGGAAGAGACCCTCACTGTGAAGTTGGCCTCATTGTTGCACC 2040

Qy   2041 GGAAGCAACGCCTGCTACATGGAACGGATGGCGTAAATGTGGGGTGGGGG 2100

Qy   2101 GGACGGGATGTGTCACAGTGGAGTGGGCCAGTTTGGGGACAATGCTGCCTGGATGAC 2160

Qy   2161 TTGCCGGACCGTGTTGATGTTGCTGTTGATGAGCTTTCTCTCAAACCGGCAAAACAGG 2220

Qy   2221 TTCGAGAAGATGAATCAGCGGCATGTACGGCACACATTCTCATTGGAT 2280

Qy   2281 TTCACGAAGGCGGGGGCTTGCTCTTCCGAGGCCGCGATCTGAGGCGCCCTCAAGACACAAGGGGA 2340
```

```
Db   2281 TTCACGAAAGGGGGGGGCTGCTTCCGAGCCGCATCGTCGAGACGGCCTCAAGACACAGGGGA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   2341 ATCTCTGAAACTAAGTTCCTGTGTCTCAGATAGAGACGGACTGCCTACCCTTGCTACAGGTT 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 ATCTCTGAAACTAAGTTCCTGTGTCTCAGATAGAGACGGACTGCCTACCCTTGCTACAGGTT 2400

Qy   2401 CGTGCCATCCTGCGCCACCTAGGGCTGGAGCAGCACGTGGACAATACGTGAAG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CGTGCCATCCTGCGCCACCTAGGGCTGGAGCAGCACGTGGACAATACGTGAAG 2460

Qy   2461 GAGGTGTGCCACTGTGGTTGCCCGGCGGCCGCGCC 2460
          |||||||||||||||||||||||||||||||||||
Db   2461 GAGGTGTGCACTGTGGTTGCCCGGCGGCCGCGCC 2520

Qy   2521 GTAGTGGACAAGATAAGAGAGAACCCGGCGCGCGCTCGACACACCCCAAAGTGACAGTGGGCGTG 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 GTAGTGGACAAGATAAGAGAGAACCCGTGGCGCTGGACAACCCCAAAGTGACAGTGGGCGTG 2580

Qy   2581 GACGGGACTCTGTATAAGCTTCANCCTCACTTTGCCAAGGTCANTGAGACGGTGAGA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GACGGGACTCTGTATAAGCTTCANCCTCACTTTGCCAAGGTCANTGAGACGGTGAGA 2640

Qy   2641 GATCTGGCTCCGAAATGTGACGTGCGTCCTTCCTTCTGTGCGACGACCAGTGGGGAGGGA 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GATCTGGCTCCGAAATGTGACGTGCGTCCTTCCTTCTGTGCGACGACCAGTGGGGAGGGA 2700

Qy   2701 GCAGCCTCTCATCACTGCCGTGGCCTGCGCCATCCGGGAGGCGGACAGCGCAGTGGGCGCGTG 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GCAGCCTCATCACTGCCGTGGCCTGCGCCATCCGGGAGGCGGACAGCGCAGTGGGCGCGTG 2700

Qy   2754 GCAGCCTCTCATCACTGCCGTGGCCTGCGCCATCCGGGAGGCGGACAGCGCAGTGGGCGAGGGA 2754
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2700 GCAGCCTCATCACTGCCGTGGCCTGCGCCATCCGGGAGGCGGACAGCGCAGTGGGCGAGGGA 2700
```

RESULT 2

AAT78599

ID   AAT78599 standard; cDNA; 2770 BP.
XX
AC   AAT78599;
XX
DT   31-OCT-1997 (first entry)
XX
DE   AS-30D tumour Type II hexokinase encoding cDNA.
XX
KW   Response element; 2-DNA; neoplasia; hexokinase II;
KW   cancer; gene therapy; diabetes; tumour; rat; ss; glycolysis;
XX
OS   Rattus rattus.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   18..2770
FT   /*tag= a
FT   /product= AS-30D_Type_II_hexokinase
FT   /note= "the last base of the stop codon is not shown"
XX
PN   WO9704104-A2.
XX
PD   06-FEB-1997.
XX
PF   12-JUL-1996; 96WO-US11673.
XX
PR   14-JUL-1995; 95US-0001199.
XX
PA   (UYJO) UNIV JOHNS HOPKINS.
XX
PI   Mathupala SP, Pedersen PL, Rempel A;
XX
DR   WPI; 1997-132643/12.
DR   P-PSDB; AAW23793.
XX
PT   New transcription regulating fragments of hexokinase II DNA contg.
PT   response element - and methods for diagnosis or treatment of
PT   neoplasias that over-express hexokinase II and for regulating
PT   glycolysis

PS   Disclosure; Fig 6; 104pp; English.
XX
CC   The present sequence represents the AS-30D tumour Type II hexokinase
CC   gene, which is a new isolated hexokinase II. DNA fragments from the
CC   promoter region (shown in AAT78598), coding sequences and introns are
CC   able to regulate transcription of a downstream open reading frame if
CC   they contain at least one response element (transcription factor
CC   binding site). A method has been produced for diagnosing a neoplasia
CC   that over-expresses hexokinase II gene. The DNA fragments may be greater than
CC   2 for the hexokinase II gene. The DNA fragments may be coupled to a
CC   reporter gene and used to screen for potential drugs that affect
CC   regulated transcription of tumour hexokinase II. Alternatively they may
CC   be coupled to a toxic gene and used to treat cells that over-express
CC   hexokinase II, such as those present in patients with cancer. They may
CC   also be used in gene therapy to treat diabetes. The DNA fragments can
CC   increase glycolysis in cells and express hexokinase. The DNA fragments
CC   protein. Probes of the DNA fragment are used in the method for
CC   diagnosing a neoplasia that over-expresses hexokinase. The new response
CC   elements are active only in tumours, not in normal cells.
XX
SQ   Sequence 2770 BP; 645 A; 666 C; 874 G; 585 T; 0 other;

```
Query Match    100.0%;    Score 2753;    DB 18;   Length 2770;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 2753;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     1 ATGATGCGCCTCGCAATAATGATCGCGCTGCTTATTCACGAGGCTCAACCAAAACCAAGTGCAG 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    18 ATGATGCGCCTCGCAATAATGATCGCGCTGCTTATTCACGAGGCTCAACCAAAACCAAGTGCAG 77

Qy    61 AAGGTTGACCAAATTTCTCTACCACATGCGGTCTTCTCAGATGAGACCCTTCTGGAGATTTCT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    78 AAGGTTGACCAAATTTCTCTACCACATGCGGTCTTCTCAGATGAGACCCTTCTGGAGATTTCT 137

Qy   121 AGGCGGTTCCGGAAGGAGAGTGGAGAAGGGCTAGGAGCTACCACGCACCCTACCAGCAGCT 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   138 AGGCGGTTCCGGAAGGAGAGTGGAGAAGGGCTAGGAGCTACCACGCACCCTACCAGCAGCT 197

Qy   181 GTGAAAAATGTTGCCTACCTTTGTGAGGTCAACTCGGATGGGGACAGAACATGGGGAGTTC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198 GTGAAAAATGTTGCCTACCTTTGTGAGGTCAACTCGGATGGGGACAGAACATGGGGAGTTC 257

Qy   241 CTGGCCTCCTGGAATCTTGGAGGAACCAACTTCCGTGTGCTCCGACTAAGGGTGACGGACAAAT 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258 CTGGCCTCCTGGAATCTTGGAGGAACCAACTTCCGTGTGCTCCGACTAAGGGTGACGGACAAAT 317

Qy   301 GGCCTCCAGAGAGTGGAGATGGAGAAACCAGAACTACACGCCACCATCCTTGAGGACATCATGCG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318 GGCCTCCAGAGAGTGGAGATGGAGAAACCAGAACTACACGCCACCATCCTTGAGGACATCATGCG 377

Qy   361 GGCAGTGGGAACCAGCCTGTTTGACCACATCGCGAATCCTGGCCAACTTCATGGACAAG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   378 GGCAGTGGGAACCAGCCTGTTTGACCACATCGCGAATCCTGGCCAACTTCATGGACAAG 437

Qy   421 CTACAAATCAAAGAAGAAGAAGCTCCCTCTCGGGTTTCACCTTTCGTTCCCCTGCCACCAG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   438 CTACAAATCAAAGAAGAAGAAGCTCCCTCTCGGGTTTCACCTTTCGTTCCCCTGCCACCAG 497

Qy   481 ACAAAACTGGATGAGACTTTTTGGTCTCGTGGACTAAGGGGTTCAAGTCCAGTGGGGTG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   498 ACAAAACTGGATGAGAGTTTTTTGGTCTCGTGGACTAAGGGGTTCAAGTCCAGTGGGGTG 557

Qy   541 GAAGGCGAGAGATGTGGTGGACCTGACCTCGATCCGGAAGGTTATCCAGCAGCAGAGGGGAC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   558 GAAGGCGAGAGATGTGGTGGACCTGACCTCGATCCGGAAGGTTATCCAGCAGCAGAGGGGAC 600

Qy   601 ATTGACATTGTGGCCGTGGTGAATGACACAGTTGGACCACATGATGACATGTGGGCTATGAT 617
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   618 ATTGACATTGTGGCCGTGGTGAATGACACAGTTGGACCACATGATGACATGTGGGCTATGAT 677

Qy   661 GATCAGAACTGCCAGATTGGTCTCATTGTGGGCACTGGCACCAACGCCTGCTACATTGAG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   678 GATCAGAACTGCCAGATTGGTCTCATTGTGGGCACTGGCACCAACGCCTGCTACATGGAG 737
```

```
Qy   721  GAAATGCCTTCATATTGACATGGTGGGAGGGAGGATGAGGGAGG  780
          ||||||||||||||||||||||||||||||||||||||||||||
Db   738  GAAATGCCGTTCATATTGACATGGTGGGAGGGAGGGAGGGCCATGTGCATCCAACATGGAG  797

Qy   781  TGGGGAGCCTTTGGGGACGACGGTACACTCAATGACATCAATGACAC  840
          |||||||||||||||||||||||||||||||||||||||||||
Db   798  TGGGGAGCCTTTGGGGACGACGGTACACTCAATGACATCAATGACACTTGACCGAGAG  857

Qy   841  ATCGACATGGGCTCGCTGAACCCTGGGAAGCAGCTGATTAGGCGGGATG  900
          |||||||||||||||||||||||||||||||||||||||||||
Db   858  ATCGACATGGGCTCGCTGAACCCTGGGAAGAAGCAGCTGATTAGGCGGGATG  917

Qy   901  TACATGGGGGAGGCTGGTCAGGGCTCAGGCTCATCCTGGTGAAGATGGGCTGT  960
          |||||||||||||||||||||||||||||||||||||||||||
Db   918  TACATGGGGGAGGCTGGTCAGGGCTCATCCTGGTGAAGATGGGCTTGTTGTTC  977

Qy   961  CAAGGGAAAACTCAGCCAGAACTCCTTACCACTCTTCGAGACCAAAAGAGTGTTCTCG  1020
          |||||||||||||||||||||||||||||||||||||||||
Db   978  CAAGGGAAAACTCAGCCAGAACTCCTTACCACTGCCTCGAGACCAAAAGAATGTCTCG  1037

Qy  1021  GATATTGAAGAGGAATAAGGATGGAAATGGAGAAGTGGGAAATGGGCTGTGTTGTTC  1080
          ||||||||||||||||||||||||||||||||||||||||
Db  1038  GATATTGAAGAGGATAAGGATGGAAATGGAGAAGGCTACCAAATCCTGATGCGCGCTGGGT  1097

Qy  1081  CTGAATCCATTGCAGGAGGATTGTGTCGGCCACCGCACCACCGCAAATCTGCCCAGAAGTGTCCACG  1140
          ||||||||||||||||||||||||||||||||||||||||
Db  1098  CTGAATCCATTGCAGGAGGATTGTGTGGGCCACCGCACCACCGCAAATCTGCCCAGAAGTGTCCACG  1157

Qy  1141  CGCTCGCGCCAGTCGTCTGTGCGCCCAGCCACCGTCTCTGTCTCGTGGGCGAATCAAAGAGAAC  1200
          ||||||||||||||||||||||||||||||||||||||||
Db  1158  CGCTCGCGCCAGTCGTCTGTGCGCAGCCACCGTCTCTGTCTCGTGGGCGAATCAAAGAGAAC  1217

Qy  1201  AAGGGCGAGGAGGCGAGCGACTTCGGCTTCGCTCACCATCGGTGTGTCGCCCGCCGGGCCGTCTCACGAAGAAACAT  1260
          ||||||||||||||||||||||||||||||||||||||||
Db  1218  AAGGGCGAGGAGGCGAGCGACTTCGGCTTCGCTCACCATCGGTGTGTTCGATGGCTCCGTCTACAAGAAACAT  1277

Qy  1261  CCCCATTTTGCCAAAGCGTCTCCATAAGGCAGTGGAGGGGTGGTGCCCGGCCGGGCCGTCTGATGTC  1320
          ||||||||||||||||||||||||||||||||||||||||
Db  1278  CCCCATTTTGCCAAGCGTCTCCATAAGGCAGTGGAGGGGTGGTGCCCGACTGTGATGTC  1337

Qy  1321  CGCTTCCTCGCGCTCTGAGGATGGCACGGGGTGCTGCTATGGTGGTACGGGCGGTGGCT  1380
          ||||||||||||||||||||||||||||||||||||||||
Db  1338  CGCTTCCTCCGCGCTCTGAGGATGGCACGGGGTGCTGCTATGGTGGTACGGGCGGTGGCT  1397

Qy  1381  TACCGTCTGGCTGACCACACCACCGGAGTCTCTGAAAGCTGAGC  1440
          ||||||||||||||||||||||||||||||||||||||||
Db  1398  TACCGTCTTCTGCTGACCACACCGGAGTCTCTGAAAGCTGAGC  1457

Qy  1441  CACGAGCACGCATGCGGGGTTCTCGGAGTGAAATGGACACAGGTGAGC  1500
          ||||||||||||||||||||||||||||||||||||||||
Db  1458  CACGAGCACGCATGCGGGGTTAAGGAGAAATGGACAGGGTCTCGAGC  1517

Qy  1501  AAGGAGACGCATGCGGTGCGCCCCCCGCGCGCCGTGTGAAGATGCTGGGGACTGGAAATGGACAGGTGAGC  1560
          ||||||||||||||||||||||||||||||||||||||||
Db  1518  AAGGAGAGCATTGCGGGGCCCCCCGAGAATGCTCCCCACTTAACGTGTGGTGCCACTGGTC  1577

Qy  1561  GATGGCACAGAGAAAAAGGAGGACTTCTTCTTGGCCCTTGGAATGTCGCCACTCCGGGTC  1620
          ||||||||||||||||||||||||||||||||||||||||
Db  1578  GATGGCACAGAGAAAAAGGAGGACTTCTTGGATCTTGGAGGAACAAACTTCCGGGTC  1637

Qy  1621  CTGCTCGGTTGCCGTGTGCGTAATGCGACCACAGCTGTGTCAATAACAGATCTAC  1680
          ||||||||||||||||||||||||||||||||||||||||
Db  1638  CTGCTCGGTTGCCGTGTGCGTAATGCGACCAGCTGTGTCAATAACAGATCTAC  1697

Qy  1681  TCCATCCCACAGAGGGTTTATTGCATTGGCATGGGGGGAAGAGCTCTTCGACCACCATTGTCCAG  1740
          ||||||||||||||||||||||||||||||||||||||||
Db  1698  TCCATCCCACAGAGGTTTATGCATTGGCATGGGGGAAGAGCTCTTTCGACCACATTGTCCAG  1757

Qy  1741  TGCATTGCGGACTTCCTCGAGTACATGGGCATGATGAAGGGCGTGTCCTGCCTTTGGGTTTC  1800
          ||||||||||||||||||||||||||||||||||||||||
Db  1758  TGCATTGGGACTTCCTCGAGTACATGGGCATGATGAAGGGCGTGTCCTGCCTTTGGGTTTC  1817


Qy  1801  ACATTCTCTTCCTTCCCTTTGCCCAGCAGAACAGCCTTAGCACCGAGAGCAGCAATCCTCCTCAAGTGGACA  1860
          ||||||||||||||||||||||||||||||||||||||||
Db  1818  ACATTCTCCTTCCTTTGCCCAGCAGAACAGCCTTAGCACCGAGAGCAGCAATCCTCCTCAAGTGGACA  1877

Qy  1861  AAGGGATTCAAAGCACATCTGGCGTCGCGAGACAGGGTGGAGTGGTGGTTGCACCTTGGTTTGG  1920
          ||||||||||||||||||||||||||||||||||||||||
Db  1878  AAGGGATTCAAAGCACATCTGGCGTCGCGAGACAGGGTGGAGTGGTGGTTGCTGAAGGAAGCG  1937

Qy  1921  ATTCACCGCGGAGGAGAGTTGACCTGGATGTGTGCCGGAGGTGGTGTGACACAGTTGGG  1980
          ||||||||||||||||||||||||||||||||||||||||
Db  1938  ATTCACCGCGGAGGAGAGTTTGACCTGGATGTGTGCCGGTGGTGTGACACAGTTGGG  1997

Qy  1981  ACTATGATGACTTGTTGGGCTACGGAGGAGAACCCTCACTGTGAAGTGGCCACC  2040
          ||||||||||||||||||||||||||||||||||||||||
Db  1998  ACTATGATGACTTGTTGGGCTACGGAGGAGAACCCTCACTGTGAAGTTGGCCACC  2057

Qy  2041  GGAAGCAAACGCCTGCTACACATGGAAGAGAAGATGCGTAATGTCGGGCTGGGCATTT  2100
          ||||||||||||||||||||||||||||||||||||||||
Db  2058  GGAAGCAAACGCCTGCTACACATGGAAGAGAAGATGCGTAATGTCGGGCTGGGCGATGAC  2117

Qy  2101  GGACGGATGTGTTTGATGTTGCTCTGTGTGGATGAGCTTTCTCTCAACAACCTGAGAG  2160
          ||||||||||||||||||||||||||||||||||||||||
Db  2118  GGACGGATGTGTGTCAACATGGAAGTTGCTCTGTGGGGCATTTGGGACAAATGGCTGC  2177

Qy  2161  TTGCGGACCGTGTTTGATGTTGCTCTGTGTGGATGAGCTTTCTCTCAACAACCTGAGAG  2220
          ||||||||||||||||||||||||||||||||||||||||
Db  2178  TTGCGGACCGTGTTTGATGTTGCAACATGGATAGTGACAACATTCTCATCGAT  2237

Qy  2221  TTCGGAGAAGATGATCAGGGGACATCGTAACTACTTGGGACGGATTGTCGCAACATTCTCATCGAT  2280
          ||||||||||||||||||||||||||||||||||||||||
Db  2238  TTCGGAGAAGATGATCAGGGGACATCGTAACTACTTGGGACGGATTGTCGCAACATTCTCATCGAT  2297

Qy  2281  TTCACGGAAGGGGGGGGCTGCTCTTCCGAGGCGCGCCGATCTCAGAGCTGCCTGCTACAGGTT  2340
          ||||||||||||||||||||||||||||||||||||||||
Db  2298  TTCACGGAAGGGGGGGGCTGCTCTTCCGAGGCGCCCGATCTCAGAAGAACAAGGGGA  2357

Qy  2341  ATCTCTGAAACTAAGTTCCTGTCTCTGAGATAGAGAGGAGGAGACTGCCTAGCGCGGTT  2400
          ||||||||||||||||||||||||||||||||||||||||
Db  2358  ATCTCTGAAACTAAGTTCCTGTCTCACGATAGAGAGGAGGAGACTGCCTAGCGCGTACAGGTT  2417

Qy  2401  CGTGCCATCCGGCCCACCTAGGCGTGGAGAGAGCACGTGGCGATGAGCAAGTCGTGAAG  2460
          ||||||||||||||||||||||||||||||||||||||||
Db  2418  CGTGCCATCCGGCCCACCTAGGCGTGGAGAGAGCACGTGGCGATGAGCATCATCGTGAAG  2477

Qy  2461  GAGGTGTGCACTGTGGTTGGCCGGGCGCGCACACAGGTCTGTGTGGGGCCGCC  2520
          ||||||||||||||||||||||||||||||||||||||||
Db  2478  GAGGTGCACTGTGGTTGCCCGGGCGCGCACACGTCTGTGTGGGGCCGCC  2537

Qy  2521  GTAGTGGACAAGATAAGGAGAACCGTGGGCGTGGACAACCCCAAAGTGACAGTGGGCGTG  2580
          ||||||||||||||||||||||||||||||||||||||||
Db  2538  GTAGTGGACAAGATAAGGAGAACCGTGGGCGTGGACAACCCCAAAGTGACAGTGGGCGTG  2597

Qy  2581  GACGGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCAATGGAAGGA  2640
          ||||||||||||||||||||||||||||||||||||||||
Db  2598  GACGGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATGAGACGGGTGAGA  2657

Qy  2641  GATCTGGCTCCGAAATGTGACGTGTCCTTCCTTGGAATCCGGGAAGGGGAGGGCAGAGATA  2700
          ||||||||||||||||||||||||||||||||||||||||
Db  2658  GATCTGGCTCCGAAATGTGACGTGTCCTTCCTTGGAATCCGGGAAGGGGAGGGCAGAGATA  2717

Qy  2701  GCAGGCTCTCATCACTGCCGCGTGGCCTGCCGCGCAATCCGGGAGGCGGCAGAGATA  2753
          ||||||||||||||||||||||||||||||||||||||||
Db  2718  GCAGCTCTCATCACTGCCGCGTGGCCTGCCGCGCAATCCGGGAGGCGGCAGAGATA  2770
```

RESULT 3
AAV00125
ID   AAV00125 standard; cDNA; 3635 BP.
XX
AC   AAV00125;
XX
DT   17-MAR-1998 (first entry)
XX

```
DE   Rat hexokinase II encoding cDNA.
XX
KW   Hexokinase; competitive enzyme; inhibitor; insulin; diabetes; B cell;
KW   glucose; glucokinase; ss;
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FH
FT   CDS             198..2951
FT                   /*tag= a
FT                   /product= Hexokinase_II
XX
PN   WO9726322-A2.
XX
PD   24-JUL-1997.
XX
PF   17-JAN-1997; 97WO-US00786.
XX
PR   19-JAN-1996; 96US-0588976.
XX
PA   (BETA-) BETAGENE INC.
PA   (UNMS ) UNIV MICHIGAN STATE.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Becker TC, Han H, Newgard CB, Wilson JE;
XX
DR   WPI; 1997-385327/35.
DR   P-PSDB; AAW37437.
XX
PT   Mammalian cell with reduced activity of low Km hexokinase - caused
PT   by competitive enzyme inhibitor, particularly an insulin secreting
PT   cell for in vitro or in vivo insulin production
XX
PS   Claim 14; Page 186-189; 254pp; English.
XX
CC   A mammalian cell has been developed which includes enough of a
CC   competitive inhibitor of low Km (Michaelis constant) hexokinase to
CC   reduce the hexokinase activity in the cell. The present sequence
CC   encodes rat hexokinase II. The cells are particularly used to produce
CC   insulin, in response to glucose or other secretagogues, either in vitro
CC   or in vivo (for treating diabetes), but may also be used to produce
CC   many other therapeutic proteins, e.g. from a gene under control of the
CC   insulin promoter and thus responsive to glucose. Inhibition of the
CC   inhibitor can be used more generally to reduce the growth rate of
CC   cells, e.g. hybridomas that produce antibodies.
XX
SQ   Sequence 3635 BP; 816 A; 875 C; 1098 G; 846 T; 0 other;

Query Match             92.6%;  Score 2550;  DB 18;  Length 3635;
Best Local Similarity   99.8%;  Pred. No. 0;  Mismatches 4;  Indels 0;  Gaps 0;
Matches  2750;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy     1 ATGATCGGCCTCGGCATAATGATCGCCTGCTTATTCACGGAGCTCAACAAAACCAAGTGCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198 ATGATCGGCCTCGGCATAATGATCGCCTGCTTATTCACGGAGCTCAACAAAACCAAGTGCAG 257

Qy    61 AAGGTTGACCAATTTCTCTACCACATGCGTCTCTCAGATGAGACCCTTCTGGAGATTTCT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258 AAGGTTGACCAATTTCTCTACCACATGCGTCTCTCAGATGAGACCCTTCTGGAGATTTCT 317

Qy   121 AGGCGGTTCCGGAAGAGAGATGGAGAAAGGGTAGGAGCTACCACGCCACCCTACAGCAGCT 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318 AGGCGGTTCCGGAAGAGAGATGGAGAAAGGGTAGGAGCTACCACGCCACCCTACAGCAGCT 377

Qy   181 GTGAAAATGTTGCCTACCTTTGTGAGGTCACACTCACGCTCAGGTCACAGAACATGGGAGTTC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   378 GTGAAAATGTTGCCTACCTTTGTGAGGTCACACTCACGCTCAGGTCACAGAACATGGGAGTTC 437

Qy   241 CTGGCTCTGGATCTTGGAGGAACCAACTTCCGTGTCGTGCTCGTGAGTAAGGGTGACGGACAAT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   438 CTGGCTCTGGATCTTGGAGGAACCAACTTCCGTGTCGTGCTCGTGAGTAAGGGTGACGGACAAT 497

Qy   301 GGCCTCAGAGAGTGGAGATGGGAGAACCAGATCTACCGCCATCCTTGAGGACATCATGCGG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   497 GGCCTCAGAGAGTGGAGATGGGAGAACCAGATCTACCGCCATCCTTGAGGACATCATGCGG 557

Db   498 GGCCTCAGAGAGTGGAGATGGGAGAACCAGATCTAGCGCCATCCTTGAGGACATCATGCGG 557
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   361 GGCAGGTGGAACCCAGCTGTTGTTGACCACACACCCGGAAATGCCTGGCCAACTCATGACAG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   558 GGCAGGTGGAACCCAGCTGTTGTTGACCACACACCCGGAAATGCCTGGCCAACTCATGACAG 617

Qy   421 CTACAAAATCAAAGAAGCAGCTCCCTTCTGGGGTTCACCTTCTCGTTCCCCTGCCACCAG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   618 CTACAAAATCAAAGAAGCAGCTCCCTTCTGGGGTTCACCTTCTCGTTCCCCTGCCACCAG 677

Qy   481 ACAAAACTGGATGAGAAGATGTTTTTTGGTCTCGTGCTGGACATCTCAGTGGGCGTG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   678 ACAAAACTGGATGAGAAGATGTTTTTTGGTCTCGTGCTGGACATCTCAGTGGGCGTG 737

Qy   541 GAAGGCCAGAGAATGTGGACCTTGGACCTTGTGGGAAGGGTTATCCGACGCGAGGGGACTTTGAC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   738 GAAGGCCAGAGAATGTGGACCTTGGACCTTGTGGGAAGGGTTATCCGACGCGAGGGGACTTTGAC 797

Qy   601 ATTGACATTGTGGCCGTGGGTGAATGACACAGTTGGGACCATGATGACTTGTGGGCTATGAT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   798 ATTGACATTGTGGCCGTGGGTGAATGACACAGTTGGGACCATGATGACTTGTGGGCTATGAT 857

Qy   661 GATCAGAACTGCGAGAGATTGGTGTCTCATTGTGTGGGGCACTGGCCAACGGCCTACATGGAG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   798 GATCAGAACTGCGAGAGATTGGTGTCTCATTGTGTGGGGCACTGGCCAACGGCCTACATGGAG 857

Qy   721 GAAATGGCTCAATATTGACATGGTGGAGGAGGAGATGAGGGGGCCGCCATGTGTCCATCAACACATGGAG 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   858 GATCAGAACTGCGAGAGATTGGTGTCTCATTGTGTGGCCAGCAACAACGGCCTACATGGAG 917

Qy   781 TGGGGAGCCTTTGGGGACGACGGTACACTCAATGACATCCGAACCGAGTTTGACCAGAG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   918 GAAATGGCTCAATATTGACATGGTGGAGGAGGAGATGAGGGGCCGCCATGTGTCCATCAACATGGAG 977

Qy   841 ATTGACACATGGGGCTGCTGTCGAACCCCTGGGAAGCAGCTGTTGTTGAGAATGAGTTAGCGGGATG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   978 TGGGGAGCCTTTGGGGACGACGGTACACTCAATGACATCCGAACCGAGTTTGACCAGAG 1037

Qy   901 TACATGGGGGAGCTGTGTCAGGTGGTCAAGGAGAGAGGCTGTTTGACGATGAGTTGTGTC 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1038 ATGGACACATGGGGCTGCGCTGTGAACCCCTGGGAAGCAGCTGTTGTTGAGAATGAGTTAGCGGGATG 1097

Qy   961 CAAGGGGAAACTCAGCCCAGAACTCCTTACCACTGGCTCCTTCGAGACCAAAAGATGTCTCG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1098 TACATGGGGGAGCTGTGTCAGGTGGTCAAGGAGAGAGGCTGTTTGACGATGAGTTGTGTC 1157

Qy  1021 GATATTGAAGAGGGATAAGGATGGAAGTCGAGAAGGCCTPACCAAATCCTGATGCGCGGGT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1158 CAAGGGGAAACTCAGCCCAGAACTCCTTACCACTGGCTCCTTCGAGACCAAAAGATGTCTCG 1217

Qy  1081 CTGAATCCATTGCAACGAGGAGGATTGTGTGGCCCAGCGCACCACGGAATTGCCAGATTGTGTCCAGG 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1218 GATATTGAAGAGGGATAAGGATGGAAGTCGAGAAGGCCTACCAAATCCTGATGCGCGGGT 1277

Qy  1141 CGCTCGGCGCAGTCTGTGTGGCCCAGCGCACCACCCGGTGGCGAGGATCAAAGAGAAC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1278 CTGAATCCATTGCAACGAGGAGGATTGTGTGGCCCAGCGCACCAGATTGTGTCCACG 1337

Qy  1201 AAGGGCGAGGAGGGCACTTCGCTCCTCCACTCGGTGTCGATGGCCTCGCCGTCTACAGCGACT 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1338 CGCTCGGCGCAGTCTGTGTGGCCCAGCGCACCACCCGGTGGCGAGAATCAAAGAGAAC 1397

Qy  1261 CCCCATTTTGCCAAGGCGTCTCCAATAAGGCCAGTGAGTGAGGCTGGTTGTCGCCGACTGTGATGTC 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1398 AAGGGCGAGGAGGGCACTTCGCTCCTCCACTCGGTGTCGATGGCCTCGCCGTCTACAGCGACT 1457

Qy  1321 CGCTTCCCTCGCGCTCTGAGGATGCGCAGGCGGCAAGGGGGCTGCTAGGTGACGAGGTGGCT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1458 CCCCATTTTGCCAAGGCGTCTCCAATAAGGCCAGTGAGTGAGGCTGGTTGTCGCCGACTGTGATGTC 1517

Qy  1381 TACCGTCTGGCTGGCCAACACCGGCCCCGCCAGAGAACCCTGGAGTCTCTGAAGCTGAGC 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1518 CGCTTCCCTCGCGCTCTGAGGATGCGCAGGCGGCAAGGGGGCTGCTAGGTGACGAGGTGGCT 1577
```

```
Db 1578 TACCGTCTGGCTGACCAAACACCGGGCCGCCGAGAAGAGCTCTGAAGCTGAGC 1637
Qy 1441 CACGCAGCAGCTTCTCGGAGGTTAAGGAGAAGAATGAAAGGGAGCAGGGTCTGAGC 1500
        |||||||||||||||||||||||||||||||||| ||||| |||||||||||||||
Db 1638 CACGCAGCAGCTTCTCGGAGGTTAAGGAGGAGAATGAAAGTGGAGCAGGGTCTGAGC 1697
Qy 1501 AAGGAGAGACCATGGGGTCGCCCCTGTGAAGATGCTGCCACTTACGTGTCACTCCA 1560
        |||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db 1698 AAGGAGAGACCATGGGGTCGCCCCTGTGAAGATGCTGCCACTTACGTGTGCCACTCCA 1757
Qy 1561 GATGGCACAGAGAAAGGAGGACTTCTTGGCCTTGGATCTTCCGGGTC 1620
        |||||||||||||||||||||||||||||||||||||||||||||||
Db 1758 GATGGCACAGAGAAAGGAGGACTTCTTGGCCTTGGATCTTCCGGGTC 1817
Qy 1621 CTGCTGCTGGTGCGTGTCGTAATGGCAAGCGGAGGGGCGTGGGAGGAGGAGGAGGCTAC 1680
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1818 CTGCTGCTGGTGCGTGTCGTAATGGCAAGCGGAGGGGCGTGGGAGGAGGAGGAGGCTAC 1877
Qy 1681 TCCATCCCACACGAGGAGGTTATGCACTGGGGAAGAGCTCTTCGACCACATTGTCCAG 1740
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1878 TCCATCCCACACGAGGAGGTTATGCATTGGGGAAGAGCTCTTCGACCACATTGTCCAG 1937
Qy 1741 TGCATTGCGGACTTCCTGGAGTACACATGGGCATGAAGGGCGTGTCCCTGCCTTTGGGTTTC 1800
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1938 TGCATTGCGGACTTCCTGGAGTACACATGGGCATGAAGGGCGTGTCCCTGCCTTTGGGTTTC 1997
Qy 1801 ACAATTCTCTTCCCTTGCCAGACAAGCAGCATCCTCCTCAAGTGGACA 1860
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1998 ACAATTCTCTTCCCTTGCCAGACAAGCAGCATCCTCCTCAAGTGGACA 2057
Qy 1861 AAGGGATTCAAAGGCATCTGGCTGCAGAGGGTGAGGAGGAAGCG 1920
        ||||||||||||||||||||||||||||||||||||||||||||
Db 1920 AAGGGATTCAAAGGCATCTGGCTGCAGAGGGTGAGGAGGAAGCG 1861 ... 
Qy 1921 ATTCACCCGCGAGGAGTTTGACCTGGATGTGGTTGCCGGTGAATGACACAGTTGGG 1980
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2058 AAGGGATTCAAAGGCATCTGGCTGCAGAGGGTGAGGAGGAAGCG 2117
Qy 1981 ACTATATGATGACTTGTGTGCTACGAAGACACCCTCACTGTGAAGTTGGCCTCATTGTTGGCACC 2040
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2178 ACTATGATGACTTGTGTGCTACGAAGACACCCTCACTGTGAAGTTGGCCTCATTGTTGGCACC 2237
Qy 2041 GGAAGCAACGCGCTGCTACATGGAAGAGGATGCGTGTAATGTGGAGCTGGTGGACGGAGGAG 2100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2238 GGAAGCAACGCCCTTGCTACATGGAAGAGGATGCGTAATGTGGAGCTGGTGGACGGAGGAG 2297
Qy 2101 GGACGGACGATGTGTGTTGTCACAACATGGGGGGAGCAATTTGGGGACAAATGGCTGCCTGGAATGAC 2160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2298 GGACGGACGATGTGTGTCAACAACATGGGGGGAGCAATTTGGGGACAAATGGCTGCCTGGATGAC 2357
Qy 2161 TTGCGGACCGCGTGTTTGATGTTGGCTGGTGGGATTGTGTGCGCAACCTGCAACATTCTCATCGAT 2220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2358 TTGCGGACCGCGTGTTTGATGTTGGCTGGTGGGATGAGCTTTCTCTCAACCCTGGCCAAACACAGAGG 2417
Qy 2221 TTCGAGAAGAATGATTCAGGGGGACATGGTACTTGGGGAGGAATTGTGCGCAACATTCTCATCGAT 2280
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2418 TTCGAGAAGAATGATCACGCGCATGTACTTGGGGAGGAATTGTGCGCAACATTCTCATCGAT 2477
Qy 2281 TTCACGAAGCGGGGGGCTGCTCTTCCGAGGGCGCCCCATCTCAGACGACAAGGGGA 2340
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2478 TTCACGAAGCGGGGGGCTGCTCTTCCGAGGGCGCCCATCTCAGACGACAAGGGGA 2537
Qy 2341 ATCTCTCGAAACTAAGTTCCTGTCTCAGATAGAGAGGCCCTAGCCTTGCTACAGGTT 2400
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2538 ATCTTTCGAAACTAAGTTCCTGTCTCAGATAGAGAGGCCCTAGCCTTGCTACAGGTT 2597
Qy 2401 CGTGCCAATCCTCCGCGCCACCTAGGGCTCGGATGACGACAACACCATCATCGTGAAG 2460
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2598 CGTGCCAATCCTCCGCGCCACCTAGGGCTCGGATGACGACAACACCATCATCGTGAAG 2657
Qy 2461 GAGGTGTTGCACTGTGGTTGCCCGGCGCCGCGCCCTCCACACGCTCTCTGTGGCCGGCCATGGCCGCC 2520
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2658 GAGGTGTTGCACTGTGGTTGCCCGGCGCTCGCACAGCTCTCTGTGGCCGGCCATGGCCGCC 2717
```

```
Qy 2521 GTAGTGGACAAGATAAGAGAGAAGAACCCGTGGGGCTGGACAAACCCAAAGTGACAGTGGGCGTG 2580
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2718 GTAGTGGACAAGATAAGAGAGAAGAACCCGTGGGGCTGGACAAACCCAAAGTGACAGTGGGCGTG 2777
Qy 2581 GACGGGACTCTGTATAAAGCTTCATCCTCACTTTGCCAAGGTCATGCATGTGAGACGGTGAGA 2640
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2778 GACGGGACTCTGTATAAAGCTTCATCCTCACTTTGCCAAGGTCATGCATGTGAGACGGTGAGA 2837
Qy 2641 GATCTGCGCTCCGAAAATGTGACGTGTGCCTTCTCTGGAATCCCAGGAACGGCAGCGCAGTGGGAAGGGA 2700
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2838 GATCTGCGCTCCGAAAATGTGACGTGTGCCTTCTCTGGAATCCCAGGAACGGCAGCGCAGTGGGAAGGGA 2897
Qy 2701 GCAGCTCTCATCACTGCCGGTGGCCTGGGAGGAGGCTGGGCAGAGATAG 2754
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db 2898 GCAGCTCTCATCACTGCCGGTGGCCTGGGAGGAGGCTGGGCAGAGATAG 2951
```

RESULT 4
AAV00086
ID   AAV00086 standard; cDNA; 3635 BP.
XX
AC   AAV00086;
XX
DT   17-MAR-1998 (first entry)
XX
DE   Rat hexokinase II encoding cDNA.
XX
KW   Yeast; trehalose-6-phosphate synthase; hexokinase; ribozyme;
KW   stimulator; insulin; glucose; diabetes; ss.
XX
OS   Rattus sp.
XX
FH   Key   Location/Qualifiers
FT   CDS   198..2951
FT   /*tag= a
FT   /product= Hexokinase_II
XX
PN   WO9726357-A1.
XX
PD   24-JUL-1997.
XX
PF   17-JAN-1997;   97WO-US00787.
XX
PR   19-JAN-1996;   96US-0588983.
XX
PA   (BETA-) BETAGENE INC.
PA   (TEXA) UNIV TEXAS SYSTEM.
XX
PI   Han H, Newgard CB, Normington KD, Thigpen AE;
XX
DR   WPI; 1997-385343/35.
DR   P-PSDB; AAW37429.
XX
PT   Mammalian cell with reduced activity of low Km hexokinase - caused
PT   by specific ribozyme or stimulator of trehalose-6-phosphate
PT   production, particularly for in vitro or in vivo insulin production
XX
PS   Claim 22; Page 186-189; 265pp; English.
XX
CC   A mammalian cell has been developed comprising an effective amount of a
CC   low Km hexokinase-inhibitor selected from: (a) an agent that stimulates
CC   production of trehalose-6-phosphate (T6P); or (b) a ribozyme specific
CC   for low Km hexokinase. The present sequence encodes rat hexokinase II.
CC   The cells are particularly used to produce insulin, in response to
CC   glucose or other secretagogues, either in vitro or in vivo (for
CC   treating diabetes), but may also be used to produce many other
CC   therapeutic proteins, e.g. from a gene under control of the insulin
CC   promoter which is therefore responsive to glucose. Growth-inhibited
CC   cells (having altered hexokinase activity) are used to produce
CC   proteins, e.g. insulin or antibodies, in vivo or in vitro. Reduction in
CC   low Km hexokinase activity provides cells in which insulin secretion is
CC   induced at glucose concentrations closer to the normal range than in

```
CC        the parent cell (nearly homeostatic secretion). Implanted cells of
CC        reduced low Km hexokinase activity are expected to survive longer in
CC        the host.
XX
SQ        Sequence 3635 BP; 816 A; 875 C; 1098 G; 846 T; 0 other;

Query Match         92.6%;  Score 2550;  DB 18;  Length 3635;
Best Local Similarity   99.8%;
Matches 2750;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;
                                   Pred. No. 0;

Qy        1 ATGATGGCCTCGCCATATGATCGGCTGCTTGTTATTCACGAGCTCAACAAAACCAAGTGCAG    60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      198 ATGATGGCCTCGCCATATGATCGGCTGCTTGTTATTCACGAGCTCAACAAAACCAAGTGCAG   257

Qy       61 AAGGTTGACCAATTTCTTACCACATCGGTCTCTCAGATGAGACCCTTCTCGGAGATTTCT   120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      258 AAGGTTGACCAATTTCTTACCACATCGGTCTCTCAGATGAGACCCTTCTCGGAGATTTCT   317

Qy      121 AGGCGGTTCCGGAAGGAGAGTGGAGAAAAGGCTAGGAGCTACCACGCACCCTACAGCAGCT   180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      318 AGGCGGTTCCGGAAGGAGAGTGGAGAAAAGGCTAGGAGCTACCACGCACCCTACAGCAGCT   377

Qy      181 GTGAAAATGTTGCCTACCTTTGTGAGGTCAACTCCGGATGGGACAGAGAACATGGGGAGTTC   240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      378 GTGAAAATGTTGCCTACCTTTGTGAGGTCAACTCCGGATGGGACAGAGAACATGGGGAGTTC   437

Qy      241 CTGGCTCTGGATCTTGGAGGAACCAACTTCGGTGCTCCGGAGTAAGGGTGAGACGACAAT   300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      438 CTGGCTCTGGATCTTGGAGGAACCAACTTCGGTGCTCCGGAGTAAGGGTGAGACGACAAT   497

Qy      301 GGCCTCCAGAGAGTGGAGATGGAGAGGAACCAGAGATCTACGCACATCCTTGAGGACATCATGGACAAG   360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      498 GGCCTCCAGAGAGTGGAGATGGAGAGGAACCAGAGATCTACGCACATCCTTGAGGACATCATGGACAAG   557

Qy      361 CTACAAATCAAAGAGAGAAAGAAGAGCTCCCTCTCGGTTTCACCTTCTCGTTCCCGTGCCACCAG   420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      558 CTACAAATCAAAGAGAGAAAGAAGAGCTCCCTCTCGGTTTCACCTTCTCGTTCCCGTGCCACCAG   617

Qy      421 ACAAAACTGGATGAGAGTTTTTTGGTCTCGTGGACTAAGGGGTTCAAGTCCAGTGGCGTG   480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      618 ACAAAACTGGATGAGAGTTTTTTGGTCTCGTGGACTAAGGGGTTCAAGTCCAGTGGCGTG   677

Qy      481 GAAGGCAGAGAGTGGTGGTGGACCTGATCCCGGAAGGTTATCCAGCGCCAGAGGGGGACTTTGAC   540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      678 GAAGGCAGAGAGTGGTGGTGGACCTGATCCCGGAAGGTTATCCAGCGCCAGAGGGGGACTTTGAC   737

Qy      541 GAAGGCAGAGAGTGGTGGGACCTGATCCCGGAAGGTTATCCAGCGCCAGAGGGGGACTTTGAC   600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      738 GAAGGCAGAGAGTGGTGGGACCTGATCCCGGAAGGTTATCCAGCGCCAGAGGGGGACTTTGAC   797

Qy      601 ATTGACATTGTGGGCGTGTCGTGAATGACACAGTTGGGCACTGATGATGACTTGTGGGTATGAT   660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      798 ATTGACATTGTGGGCCGTGTCGTGAATGACACAGTTGGGCACTGATGATGACTTGTGGGTATGAT   857

Qy      661 GATCAGAAACTGCGAGAATTGGTCTCTCATTGTGTGCACTGGCCACTGGCCCGCGCAACCGCTGCTGCTGCACATGGAG   720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      858 GATCAGAAACTGCGAGAATTGGTCTCATTGTGTGCACTGGCCACTGGCCCGCGCAACCGCTGCTGCTGCACATGGAG   917

Qy      721 GAAATGCCTCATATTGACATGGTGGGACGGAGGGAGGGGGCCATGGTCGCATCAACATGGAG   780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      918 GAAATGCCTCATATTGACATGGTGGGACGGAGGGAGGGGGCCATGGTCGCATCAACATGGAG   977

Qy      781 TGGGGAGCCTTTGGGGACGGAGCTACACTGCAACACTTGCCATGAGTGGAATCGGATG   840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      978 TGGGGAGCCTTTGGGGACGGAGCTACACTGCAACACTTGCCATGAGTGGAATCGGATG   1037

Qy      841 ATCGACAATGGGCTGCTGCTGAAGGAAGCAGCTGTTTGAGAAGATGTTAGCGGGGATG   900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1038 ATCGACAATGGGCTGCTGCTGAAGGAAGCAGCTGTTTGAGAAGATGTTAGCGGGGATG   1097

Qy      901 TACATGGGGAGCTGGTCAGGTCACTCCTTGTGAAGAATGGGGCCAAGGCAGAGCTGTTGTTC   960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1098 TACATGGGGAGCTGGTCAGGTCACTCCTTGTGAAGAATGGGGCCAAGGCAGAGCTGTTGTTC   1157

Qy      961 CAAGGGAAACTCAGCCCAGAAACTCCTTACCACTGGCCTCCTTCGACCCAAAGATGTCTCG   1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1158 CAAGGGAAACTCAGCCCAGAAACTCCTTACCACTGGCCTCCTTCGACCCAAAGATGTCTCG   1217

Qy     1021 GATATTGAAGGAGGATTAAGGATGGAAATCAGAATGGAGAAGGCCTACCAGTTGTTGTCCAG   1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1218 GATATTGAAGGAGGATTAAGGATGGAAATCAGAATGGAGAAGGCCTACCAGTTGTTGTCCAG   1277

Qy     1081 CTGAATCCATTGCAGGAGGATTGTGTGGCCAGCCAGCCACCACCGAATCGCCAGATTCAAAGGAAC   1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1278 CTGAATCCATTGCAGGAGGATTGTGTGGCCAGCCAGCCACCACCGAATCGCCAGATTCAAAGGAAC   1337

Qy     1141 CGCTCGGCCCAGTGCTTGTTGGGCAGCCAGCCACCCCTGGCCACCATGGGCCGGGTGCCAGCC   1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1338 CGCTCGGCCCAGTGCTTGTTGGGCAGCCAGCCACCCCTGGCCACCATGGGCCGGGTGCCAGCC   1397

Qy     1201 AAGGGGCGGAGGAGGGAGACTTCGCGTGTCGTCGATGGTCCGATTCCTACAAGAAACAT   1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1398 AAGGGGCGGAGGAGGGAGACTTCGCGTGTCGTCGATGGTCCGATTCCTACAAGAAACAT   1457

Qy     1261 CCCCATTTTGCCAAGGGTGTCCCAACATGCAGTGAGGAGGAGGCTGAGGAGCTGGTC   1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1458 CCCCATTTTGCCAAGGGTGTCTCCATAAGGCAGTGAGGAGGAGGCTGAGGAGCTGGTC   1517

Qy     1321 CGCTTCCTTCCCGTCTCGAGGGTTAAGGAGGAATGAAGAGAATGGAAGTCTGATGGGTGGGCT   1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1518 CGCTTCCTTCCCGTCTCGAGGGTTAAGGAGGAATGAAGAGAATGGAAGTGGAACGACCCTGAGTGTCTCTGAAGCTGAGC   1577

Qy     1381 TACCGTCGTGTGGCTCGAACCACACCGGGCCCCGGCCAAGGACCCTGCTGTATGTGGACGTCTGAAGC   1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1578 TACCGTCGTGTGGCTCGAACCACACCGGGCCCCGGCCAAGGCAGTGAAGTGTCTGAAGCTGGTGAGC   1637

Qy     1441 CACGAGCAGCTTCTGACCAAACACCGGGCCCCGGCCAAGGACCCTGCTGTATGTGGACGTCTGAAGC   1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1638 CACGAGCAGCTTCTGACCAAACACCGGGCCCCGGCCAAGGCAGTGAAGTGTCTGAAGCTGGTGAGC   1697

Qy     1501 AAGGAGACGCCATGCGGTCGCGCCCGTCGTGAAAGATGCTGCCCACTTAGGTGTGTGCCACTCCA   1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1698 AAGGAGACGCCATGCGGTCGCGCCCCGTCGTGAAAGATGCTGCCCACTTAGGTGTGTGCCACTCCA   1757

Qy     1561 GATGGCACAGAGAAGAAAGGAGACTTCTTGGGCCTTGGATCTTGGAGGAACAAAACTTCCGGGTC   1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1758 GATGGCACAGAGAAGAAAGGAGACTTCTTGGGCCTTGGATCTTGGAGGAACAAAACTTCCGGGTC   1817

Qy     1621 CTGCTGGTGCGTGTGCGTAATGGCAAGCGGAGCGGAGGGGCCGTTGAGATGCCATAACAAGATCTAC   1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1818 CTGCTGGTGCGTGTGCGTAATGGCAAGCGGAGGGGGCCGTTGAGATGCCATAACAAGATCTAC   1877

Qy     1681 TCCATCCCACAGGAGGTTATGCATGGCACTGGGAAGAGCTTCGACCACATTGTCCAG   1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1878 TCCATCCCACAGGAGGTTATGCATGGCACTGGGAAGAGCTTCGACCACATTGTCCAG   1937

Qy     1741 TGCATTGCGGACTTCCTCGAGTACATGGGCATGGCGATGGAAGGAGCCGTGTCCCTGCCC   1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1938 TGCATTGCGGACTTCCTCGAGTACATGGGCATGGCGATGGAAGGAGCCGTGTCCCTGCCCTTTGGGTTTC   1997

Qy     1801 ACAATTCTCCTTCCCTTGCCAGCAGAACACAGCCTAGACACAGAACACTCCTCCAAGTGGACA   1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1998 ACAATTCTCCTTCCCTTGCCAGCAGAACACAGCCTAGACACAGAACACTCCTCCAAGTGGACA   2057

Qy     1861 AAGGGATTCAAAGGCACTCTGGCTCCGAGGGTGAGGATGTTGGTCACCTGGGAAGAGCG   1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2058 AAGGGATTCAAAGGCATCTGGCTCCGAGGGTGAGGATGTTGGTCACCTGGGAAGAGCG   2117

Qy     1921 ATTCACCCGGCGACGAGGAGAGTTTGACCTGGATGTGGTTGCGGTGAATGACAGTGTTG   1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2118 ATTCACCCGGCGACGAGGAGAGTTTGACCTGGATGTGGTTGCGGTGAATGACAGTGTTGGG   2177

Qy     1981 ACTATGATGACTTGTGGCTACCGAAGAGACCCTCACTGTGTGAAGTTGGCACC   2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2178 ACTATGATGACTTGTGGCTACCGAAGAGACCCTCACTGTGTGAAGTTGGCACC   2237
```

```
Qy  2041 GGAAGCAACGCCCTGCTACATGGACAGAGAGATGCGTAATGTGGAGCTGGTGGACGGAGGAGGAG 2100
         ||||||||||||||||||||||||| ||||||||| |||||||||||||||||||||||||||||
Db  2238 GGAAGCAACGCCCTGCTACATGGAAGAGAGATCCGTAATGTGGAGCTGGTGGACGGAGGAGGAG 2297

Qy  2101 GGACGGGGATGTTGTGTCAACACATGGGGAGTGGGGAGCATTTGGGACACATGGCTGCCTGGATGAC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2298 GGACGGGATGTGTGTCAACACATGGGGACCATTTGGGGACAATGGCTCGCCTGGATGAC 2357

Qy  2161 TTGCGGGACCCGTGTTTGATGTTGGTTGCTTTGTCTCTCAAGCCCTGCCAAACAGAGG 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2358 TTGCGGGACCCGTGTTTGATGTTGGTTGCTTTCTCTCAAGCCCTGCCAAACAGAGG 2417

Qy  2221 TTCGAGAAGATGATCAGCGGCCATGTACTTGGGGAGAGAGATTGTGCGCAACATTTCTCATCGAT 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2418 TTCGAGAAGATGATCAGCGGCCATGTACTTGGGGAGAGATTGTGCGCAACATTTCTCATCGAT 2477

Qy  2281 TTCACAAGCGCGGGGGCTGCTCTTCCGAGGCCGCCATCTCAGACAACAAGGGGA 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2478 TTCACGAAGCGGGGGCTGCTCTTCCGAGGCCGCCATCTCAGACAACAAGGGGA 2537

Qy  2341 ATCTCTGCAAACTAAGTTCCTGTTCTCACATAGAGAGCGCACTGCCTAGCCCTGCTACAGGTT 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2538 ATCTTTGCAAACTAAGTTCCTGTTCTCAGATAGAGAGCGCACTGCCTAGCCCTGCTACAGGTT 2597

Qy  2401 CGTGCCATCCTGCGCCACCTAGGGCGTGGAGAGCCACGTGGGCAGCACCATCATCGTGAAG 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2598 CGTGCCATCCTCGCCACCTAGGGCGTGGGGAGCACCATCATCATCGTGAAG 2657

Qy  2461 GAGGTGTGCACTGGGTTGGTTGCCCGGGGCCGGCTCAGCTGTGTGCGCACACGCACAGTGGGCGG 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2658 GAGGTGTGCACTGTGGTTTGCCCGGGCCCGGCTCAGCTCGTGTGCGCACACGCACAGTGGGCGCC 2717

Qy  2521 GTAGTGTGGACAAGATAAGAGAGAGAGAACCGTGGGGTGGACAACCCAAAGTGACAGTGGGCGTG 2580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2718 GTAGTGGACAAGATAAGAGAGAAACCGTGGGGCTGGACAACCTCAAAGTGACAGTGGGCGTG 2777

Qy  2581 GACGGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCAATGCATGCATGAGACGGTGAGA 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2778 GACGGGGACTCTGTTATAAGCTTCATCCTCACTTTGCCAAGGTCAATGCATGCATGAGACGGTGAGA 2837

Qy  2641 GATCTGGCTCCGAAAATGTCACGTGTCTCCTTCCTGGAATCCGAGGACGGCGCAGTGGGGAAGGGA 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2838 GATCTGGCTCCGAAAATGTGAGCGTGTCCTTCCTGGAATCCGAGGACGGCGCAGTGGGGAAGGGA 2897

Qy  2701 GCAGCTCTCATCACTGCCGTGGGCCTGCCCGAGGCTGGGCAGAGATAG 2754
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  2898 GCAGCTCTCATCACTGCCGTGGGCCATCCGGGAGGCTGGGCAGAGATAG 2951
```

```
RESULT 5
ID   AAT80522 standard; cDNA; 299 BP.
XX
AC   AAT80522;
XX
DT   31-OCT-1997  (first entry)
XX
DE   AS-30D tumour Type II hexokinase probe.
XX
KW   Response element; Z-DNA; neoplasia; hexokinase II; glycolysis;
KW   cancer; gene therapy; diabetes; tumour; rat; PCR;
KW   polymerase chain reaction; ss.
OS   Mus musculus.
XX
PN   WO9704104-A2.
XX
PD   06-FEB-1997.
XX
PF   12-JUL-1996;  96WO-US11673;
XX
PR   14-JUL-1995;  95US-0001199.
XX
PA   (UJJO ) UNIV JOHNS HOPKINS.
XX
PI   Mathupala SP, Pedersen PL, Rempel A;
XX
DR   WPI; 1997-132643/12.
XX
PT   New transcription regulating fragments of hexokinase II DNA contg.
PT   response element - and methods for diagnosis or treatment of
PT   neoplasias that over-express hexokinase II and for regulating
PT   glycolysis
XX
PS   Example 1; Fig 7; 104pp; English.
```

The present sequence represents the cDNA of the PCR-generated probe used for the isolation of the promoter for Type II tumour hexokinase from the AS-30D genomic library. AS-30D tumour Type II hexokinase is a new isolated hexokinase II. DNA fragments from the promoter region (shown in AAT78598), coding sequences and introns are able to regulate transcription of a downstream open reading frame if they contain at least one response element (transcription factor binding site). A method has been produced for diagnosing a neoplasia that over-expresses hexokinase by detecting a copy number greater than a 2 for the hexokinase II gene. The DNA fragments may be coupled to a reporter gene and used to screen for potential drugs that affect regulated transcription of tumour hexokinase II. Alternatively they may be coupled to a toxic gene and used to treat cells that over-express hexokinase II, such as those present in patients with cancer. They may also be used in gene therapy to treat diabetes. The DNA fragments can increase glycolysis in cells and express homologous or heterologous protein. Probes of the DNA fragment are used in the method for diagnosing a neoplasia that over-expresses hexokinase. The new response elements are active only in tumours, not in normal cells.

```
SQ   Sequence 299 BP; 59 A; 96 C; 76 G; 68 T; 0 other;
```

```
Query Match             3.7%;  Score 102;  DB 18;  Length 299;
Best Local Similarity   100.0%;
Matches  102;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 ATGATCGGCCTCGCCATATGATCGCGCCTGCTTATTCACGGAGCGCTCAACCAAAACCAAGTGCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198 ATGATCGGCCTCGCCATATGATCGCGCCTGCTTATTCACGGAGCGCTCAACCAAAACCAAGTGCAG 257

Qy    61 AAGGTTGACCAAATTTCTCTACCACACATGCGTCTCTCGAGATGAG 102
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   258 AAGGTTGACCAAATTTCTCTACCACACATGCGTCTCTCGAGATGAG 299
```

```
RESULT 6
ID   AAT78598 standard; DNA; 5150 BP.
XX
AC   AAT78598;
XX
DT   30-OCT-1997  (first entry)
XX
DE   AS-30D tumour Type II hexokinase 4.3kbp proximal promoter region.
XX
KW   Response element; Z-DNA; neoplasia; hexokinase II; glycolysis;
KW   cancer; gene therapy; diabetes; tumour; rat; ds.
XX
OS   Rattus rattus.
XX
FH   Key            Location/Qualifiers
FH
FT   protein_bind   58..63
FT                  /*tag= a
FT                  /label= Myb
FT                  /note= "Transcription factor binding site"
FT   protein_bind   118..126
FT                  /*tag= b
FT                  /label= p53
FT                  /note= "Transcription factor binding site"
FT   protein_bind   171..180
FT                  /note= "Transcription factor binding site"
```

```
                    /tissue_type="leukemia"
                    /note="Inverse PCR method
                    (http://genome2.ncifcrf.gov/RTCGD)"
BASE COUNT      109 a     127 c     116 g     114 t
ORIGIN

Query Match           2.0%;     Score 55;   DB 17;  Length 466;
Best Local Similarity 100.0%;   Pred. No. 1.2e-16;
Matches   55;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAG 55
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     55 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAG 1


RESULT 7
AW823261
LOCUS       AW823261              518 bp    mRNA      linear  EST 17-MAY-2000
DEFINITION  ur71a09.y1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:3155704 5'
            similar to SW:HXK2_MOUSE O08528 HEXOKINASE TYPE II ;, mRNA
            sequence.
ACCESSION   AW823261
VERSION     AW823261.1  GI:7916338
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 518)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml
            MGI:1058460
            Seq primer: -40RP from Gibco
            High quality sequence stop: 418.
FEATURES             Location/Qualifiers
     source          1..518
                     /organism="Mus musculus"
                     /strain="129.C57BL/6J,FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3155704"
                     /clone_lib="NCI_CGAP_Mam3"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="10 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                     Reference for transgenic model: Xu et al., Nature Genetics
                     22, 37-43 (1999)."
BASE COUNT      125 a     133 c     142 g     118 t
ORIGIN

Query Match           1.9%;     Score 51;   DB 10;  Length 518;
Best Local Similarity 100.0%;   Pred. No. 1.4e-14;
Matches   51;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     13 CATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAGAAG 63
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    468 CATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAGAAG 518
```

RESULT 4
BQ951609
```
LOCUS       BQ951609                924 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8931340 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6468190
            5', mRNA sequence.
ACCESSION   BQ951609
VERSION     BQ951609.1  GI:22367087
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 924)
AUTHORS     NIH-MGC http://mgc.ncl.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAMI3995 row: e column: 23
            High quality sequence stop: 606.
FEATURES             Location/Qualifiers
     source          1..924
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6468190"
                     /clone_lib="NIH_MGC_94"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      203 a    224 c    271 g    222 t
ORIGIN

Query Match             3.0%;  Score 83;  DB 14;  Length 924;
Best Local Similarity  100.0%;  Pred. No. 5.4e-31;
Matches   83;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   2620  GTCATGCATGAGACGGTGAGAGATCTGGCTCCGAAATGTGACGTGCCTTCCTGGAATCC 2679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    151  GTCATGCATGAGACGGTGAGAGATCTGGCTCCGAAATGTGACGTGCCTTCCTGGAATCC 210

Qy   2680  GAGGACGGCAGTGGGAAGGGAGC 2702
           |||||||||||||||||||||||
Db    211  GAGGACGGCAGTGGGAAGGGAGC 233
```

RESULT 5
BI738353
```
LOCUS       BI738353                754 bp    mRNA    linear   EST 20-SEP-2001
DEFINITION  603611115F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5368490 5',
            mRNA sequence.
ACCESSION   BI738353
VERSION     BI738353.1  GI:15715366
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 754)
AUTHORS     NIH-MGC http://mgc.ncl.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAMI1937 row: m column: 03
            High quality sequence start: 4
            High quality sequence stop: 745.
FEATURES             Location/Qualifiers
     source          1..754
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5368490"
                     /clone_lib="NIH_MGC_94"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      181 a    190 c    219 g    164 t
ORIGIN

Query Match             2.6%;  Score 71;  DB 13;  Length 754;
Best Local Similarity  100.0%;  Pred. No. 7.8e-25;
Matches   71;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1  ATGATCGCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   464  ATGATCGCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG 523

Qy    61  AAGGTTGACCA 71
          |||||||||||
Db   524  AAGGTTGACCA 534
```

RESULT 6
BH859067/c
```
LOCUS       BH859067                466 bp    DNA    linear   GSS 08-JUL-2002
DEFINITION  S5_319d_s6 Mouse Retroviral Tagged Cancer Gene Database Mus
            musculus genomic clone S5_319d, DNA sequence.
ACCESSION   BH859067
VERSION     BH859067.1  GI:21709888
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 466)
AUTHORS     Suzuki,T., Shen,H., Akagi,K., Morse,H.C., Malley,J.D., Naiman,D.Q.,
            Jenkins,N.A. and Copeland,N.G.
TITLE       Retroviral tagging provides a potent cancer gene discovery tool in
            the post-genome-sequence era
JOURNAL     Nat. Genet., (2002) In press
COMMENT     Contact: Copeland NG
            Mouse Cancer Genetics Program
            National Cancer Institute
            Bldg. 539, Rm. 229, Frederick, MD 21702-1201, USA
            Tel: 301 846 1260
            Fax: 301 846 6666
            Email: copeland@ncifcrf.gov
            Class: PCR with specific primers.
FEATURES             Location/Qualifiers
     source          1..466
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="S5_319d"
                     /clone_lib="Mouse Retroviral Tagged Cancer Gene Database"
                     /sex="female"
```

```
Db    457  TTGCGGACTTCCTGGAGTACATGGCATGGCAGTGTGTCCCTGCCTTGGGTTTCACAT  398
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1805  TCTCCTTCCCTTGCCAGCAGAGACAGCCTAGACCAGAGCATCCTCCTCAAGTGGACAAAGG  1864
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    397  TCTCCTTCCCTTGCCAGCAGAACAGCCTAGACCAGAGCATCCTCCTCAAGTGGACAAAGG  338
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1865  GATTCAAGGCATCTGGCTCGCGAGGTGAGGATGTGGTCACCTTGCTGAAGGAAGCGATTC  1924
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    337  GATTCAAGGCATCTGGCTCGCGAGGTGAGGATGTGGTCACCTTGCTGAAGGAAGCGATTC  278
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1925  ACCGGCGAGGAGTTGACCTGGATGTGGTTGCCGTGGTGAATGACACAGTTGGGACTA  1984
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    277  ACCGGCGAGGAGTTGACCTGGATGTGGTTGCCGTGGTGAATGACACAGTTGGGACTA  218
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1985  TGATGACTTGTGGCTACGAAGACCCTCACTGTGAAGTTGGCCTCATTGTTGGCACCGGAA  2044
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    217  TGATGACTTGTGGCTACGAAGACCCTCACTGTGAAGTTGGCCTCATTGTTGGCACCGGAA  158
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2045  GCAACGCCTGCTACTACATGGAAGAGATGCGTAATGTGGAGCTGGTGGACG  2092
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    157  GCAACGCCTGCTACTACATGGAAGAGATGCGTAATGTGGAGCTGGTGGACG  110
```

RESULT 2
AA574507
LOCUS   AA574507   444 bp   mRNA   linear   EST 02-SEP-1997
DEFINITION   vm29b05.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA
             clone IMAGE:991569 5' similar to gb:M75126 HEXOKINASE, TYPE I
             (HUMAN); gb:J05277 Mouse hexokinase mRNA, complete cds (MOUSE);,
             mRNA sequence.
ACCESSION   AA574507
VERSION   AA574507.1 GI:2349133
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 444)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
             Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
             Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
             Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
             Waterston,R.
  TITLE   The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT   Contact: Marra M/Mouse EST Project
             WashU-HHMI Mouse EST Project
             Washington University School of MedicineP
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:563849
             High quality sequence stop: 411.
FEATURES   Location/Qualifiers
     source   1..444
                     /organism="Mus musculus"
                     /strain="C57BL/6J x DBA/2J F1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:991569"
                     /clone_lib="Knowles Solter mouse blastocyst B1"
                     /tissue_type="blastocyst"
                     /dev_stage="embryo (pre-implantation)"
                     /lab_host="DH10B"
                     /note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally from mRNA prepared
                     from 800 blastocysts.   Primer: SalI(dT):
                     5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTTTT-3'. cDNAs were
                     cloned into the NotI/SalI sites of a pSPORT vector (Life
                     Technologies). Two different size selections: B1 (larger
```
Db                                    95 a        inserts) and B3."
                                                  117 c         145 g          87 t
BASE COUNT   
ORIGIN
Query Match                          3.0%;      Score 83;   DB 9;   Length 444;
Best Local Similarity   100.0%;   Pred. No. 4.4e-31;
Matches    83;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy   2620  GTCATGCATGAGACGGTGAGAGATCGGCTCCGAAATGTGACGTGCCTTCCTGGAATCC  2679
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342  GTCATGCATGAGACGGTGAGAGATCGGCTCCGAAATGTGACGTGCCTTCCTGGAATCC  401

Qy   2680  GAGGACGGCAGTGGGAAGGGAGC  2702
           |||||||||||||||||
Db    402  GAGGACGGCAGTGGGAAGGGAGC  424

RESULT 3
BG295479
LOCUS        BG295479                   746 bp    mRNA        linear     EST 21-FEB-2001
DEFINITION   602928O8F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4504920 5',
             mRNA sequence.
ACCESSION    BG295479
VERSION      BG295479.1  GI:13057155
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 746)
  AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: The Cepko Laboratory
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLAM10377  row: o  column: 01
             High quality sequence stop: 736.
FEATURES             Location/Qualifiers
     source          1..746
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4504920"
                     /clone_lib="NIH_MGC_94"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
```
BASE COUNT   166 a   191 c   225 g   164 t
ORIGIN
Query Match   3.0%;   Score 83;   DB 12;   Length 746;
Best Local Similarity   100.0%;   Pred. No. 5.1e-31;
Matches   83;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   2620 GTCATGCATGAGACGGTGAGAGATCGGCTCCGAAATGTGACGTGCCTTCCTGGAATCC 2679
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GTCATGCATGAGACGGTGAGAGATCGGCTCCGAAATGTGACGTGCCTTCCTGGAATCC 487

Qy   2680 GAGGACGGCAGTGGGAAGGGAGC 2702
           |||||||||||||||||
Db   488 GAGGACGGCAGTGGGAAGGGAGC 510
```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2003, 13:32:46.; Search time 3815 Seconds
                 (without alignments)
                 11691.316 Million cell updates/sec

Title: US-09-808-743a-1
Perfect score: 2754
Sequence: 1 atgatcgcctcgcatgat.........gggaggctggcagagatag 2754

Scoring table: OLIGO_NUC
                 Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size :   50

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   EST:*
                 1: em_estba:*
                 2: em_esthum:*
                 3: em_estin:*
                 4: em_estmu:*
                 5: em_estov:*
                 6: em_estpl:*
                 7: em_estro:*
                 8: em_htc:*
                 9: gb_est1:*
                 10: gb_est2:*
                 11: gb_htc:*
                 12: gb_est3:*
                 13: gb_est4:*
                 14: gb_est5:*
                 15: em_estfun:*
                 16: em_estom:*
                 17: gb_gss:*
                 18: em_gss_hum:*
                 19: em_gss_inv:*
                 20: em_gss_pln:*
                 21: em_gss_vrt:*
                 22: em_gss_fun:*
                 23: em_gss_mam:*
                 24: em_gss_mus:*
                 25: em_gss_other:*
                 26: em_gss_pro:*
                 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query Match	Query Length	DB	ID	Description
c	1	348	12.6	521	10	BE108528	BE108528 UI-R-CA0-
	2	83	3.0	444	9	AA574507	AA574507 vm29b05.r
	3	83	3.0	746	12	BG295479	BG295479 602392808
	4	83	3.0	924	14	BQ951609	BQ951609 AGENCOURT
	5	71	2.6	754	13	BI738353	BI738353 603361115
c	6	55	2.0	466	17	BH859067	BH859067 S5_319d_s

ALIGNMENTS

| 7 | 51 | 1.9 | 518 | 10 | AW823261 |
| | | | | | AW823261 ur71a09.y |

RESULT 1
BE108528/c
LOCUS   BE108528   521 bp   mRNA   linear   EST 13-JUN-2000
DEFINITION   UI-R-CA0-axf-c-10-0-UI.s1 UI-R-CA0 Rattus norvegicus cDNA clone
             UI-R-CA0-axf-c-10-0-UI 3', mRNA sequence.
ACCESSION   BE108528
VERSION   BE108528.1 GI:8500633
KEYWORDS   EST.
SOURCE   Norway rat.
ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE   1 (bases 1 to 521)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE   Normalization and subtraction: two approaches to facilitate gene
             discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
COMMENT   Contact: Soares, MB
             Program for Rat Gene Discovery and Mapping
             University of Iowa
             451 Eckstein Medical Research Building Iowa City, IA 52242, USA
             Tel: 319 335 8250
             Fax: 319 335 9565
             Email: msoares@blue.weeg.uiowa.edu
             The sequence contained an oligo-dT track that was present in the
             oligonucleotide that was used to prime the synthesis of first
             strand cDNA and therefore this may represent a bonafide NotI site
             tail. The sequence tag present in the cDNA between the NotI site
             and the oligo-dT track served to identify it as a clone from the
             normalized hypothalamus library cDNA library Preparation: M.B.
             Soares Lab Clone distribution: clones will be available through
             Research Genetics (www.resgen.com) The following repetitive
             elements were found in this cDNA sequence: 1-32,
             >POLY_A#Simple_repeat
             Seq primer: M13 Forward
             POLYA=Yes.
FEATURES   Location/Qualifiers
     source   1..521
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="UI-R-CA0-axf-c-10-0-UI"
                     /clone_lib="UI-R-CA0"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CA0
                     library is a subtracted library derived from the following
                     tissues: thalamus, cerebellum, hypothalamus, medulla, pons
                     , midbrain, cerebral cortex, corpus striatum, testis, and
                     hippocampus. For a detailed description of the library
                     from which this clone was derived, please visit our web
                     site at ratest.eng.uiowa.edu. The subtraction has been
                     previously described in (Bonaldo, Lennon and Soares,
                     Genome Research 6:791-806, 1996)
                     TAG_LIB=UI-R-CA0
                     TAG_TISSUE=hypothalamus
                     TAG_SEQ=GATGC"
BASE COUNT   114 a   157 c   112 g   138 t
ORIGIN

Query Match   12.6%;   Score 348; DB 10;   Length 521;
Best Local Similarity 100.0%;   Pred. No. 1.3e-167;
Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1745 TTGCGGACTTCCTGGAGTACATGGCATGAAGGGCGTGTCCCTGCCTTTGGGTTTCACAT 1804

```
Qy   901  TACATGGGGAGCTGGTCAGGCTCATCCTGGTGAAGATGGCCAAGGCGAGCTGTTGTTC   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TACATGGGGAGCTGGTCAGGCTCATCCTGGTGAAGATGGCCAAGGCGAGCTGTTGTTC   960

Qy   961  CAAGGGAAACTCAGCCCAGAACTCCTTACCACTGGCTCCTTCGAGACCAAAGATGTCTCG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  CAAGGGAAACTCAGCCCAGAACTCCTTACCACTGGCTCCTTCGAGACCAAAGATGTCTCG  1020

Qy  1021  GATATTGAAGAGGATAAGGATCGAGAAGGCCTACCAAATCCTGATCGCGCCTGGGT     1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  GATATTGAAGAGGATAAGGATCGAGAAGGCCTACCAAATCCTGATCGCGCCTGGGT     1080

Qy  1081  CTGAATCCATTGCAGGAGGATTGTGTGGCCACGCACCGGAATCTGCCAGATTGTGTCCACG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  CTGAATCCATTGCAGGAGGATTGTGTGGCCACGCACCGGAATCTGCCAGATTGTGTCCACG  1140

Qy  1141  CGCTGGCGCCAGTCGTGCCGAGCCACCGGTCCGGTGGGCGAATCAAAGAGAAC        1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  CGCTGGCGCCAGTCGTGCCGAGCCACCGGTCCGGTGGGCGAATCAAAGAGAAC        1200

Qy  1201  AAGGGCGAGGAGGCGACTTCGCTCCACCATCGGTGTCGATGGCTCCGTCTACAAGAAACAT  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  AAGGGCGAGGAGGCGACTTCGCTCCACCATCGGTGTCGATGGCTCCGTCTACAAGAAACAT  1260

Qy  1261  CCCCATTTTGCCAAGCGTCTCCATAAGGCAGTGAGGAGGCTGGTGCCCGACTGTGATGTC   1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  CCCCATTTTGCCAAGCGTCTCCATAAGGCAGTGAGGAGGCTGGTGCCCGACTGTGATGTC   1320

Qy  1321  CGCTTCCTCCGCTCTGAGGATGGCAGCGGCAAGGGCGCCCGCCGCCTGTGCTGTCTGAAG   1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CGCTTCCTCCGCTCTGAGGATGGCAGCGGCAAGGGCGCCCGCCGCCTGTGCTGTCTGAAG   1380

Qy  1381  TACCGTCTGGCTGACCAACACCGGCCCCGCCAGAAGACCCTGGAGTCTCTGAAGCTGAGC   1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  TACCGTCTGGCTGACCAACACCGGCCCCGCCAGAAGACCCTGGAGTCTCTGAAGCTGAGC   1440

Qy  1441  CACGAGCAGCAGTTCTTGGAGTTAAGGAGTTAAGGAGAAGTGAAATGGAGCAGGGTCTGAGC   1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  CACGAGCAGCAGTTCTTGGAGTTAAGGAGTTAAGGAGAAGTGAAATGGAGCAGGGTCTGAGC   1500

Qy  1501  AAGGAGACCGCCATGTGCGGTTATGCTGCCCACTTACGTGTGTGCCACTCCA          1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  AAGGAGACCGCCATGTGCGGTTATGCTGCCCACTTACGTGTGTGCCACTCCA          1560

Qy  1561  GATGGCACAGAGAAAGGAGACTTCTTGGACTCTTGGAGGGAACAAACTTCCGGGTC      1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  GATGGCACAGAGAAAGGAGACTTCTTGGACTCTTGGAGGGAACAAACTTCCGGGTC      1620

Qy  1621  CTGCTGGTGCGTGTGCGTAATGGCAAGCGGAGGGCGTGGAGATGGCATAACAAGAATCTAC  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CTGCTGGTGCGTGTGCGTAATGGCAAGCGGAGGGCGTGGAGATGGCATAACAAGAATCTAC  1680

Qy  1681  TCCATCCCACAGGAGGTTATGCATGGCACTGGGGAAGAGCTCTTCGACCACATTGTCCAG   1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TCCATCCCACAGGAGGTTATGCATGGCACTGGGGAAGAGCTCTTCGACCACATTGTCCAG   1740

Qy  1741  TGCATTGCGGACTTCCTGGAGTACATGGGCATGGGCGTGTCCTTGCCTTTGGGTTTC      1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGCATTGCGGACTTCCTGGAGTACATGGGCATGGGCGTGTCCTTGCCTTTGGGTTTC      1800

Qy  1801  ACATTCTCCTTCCCTTGCCAGAACACCTAGACCAGAGCAATCCTCCTCAAGTGACA      1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  ACATTCTCCTTCCCTTGCCAGAACACCTAGACCAGAGCAATCCTCCTCAAGTGACA      1860

Qy  1861  AAGGGATTCAAGGCACTGGCTCCGGAGGGTGAGATGGTGTCACCTTGCTGAAGGAAGCG    1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  AAGGGATTCAAGGCACTGGCTCCGGAGGGTGAGATGGTGTCACCTTGCTGAAGGAAGCG    1920

Qy  1921  ATTCACCGGCGAGGAGTTTGACCTGGTGGTTGCCGTGGTGAATGACACCAGTTGGG      1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  ATTCACCGGCGAGGAGTTTGACCTGGTGGTTGCCGTGGTGAATGACACCAGTTGGG      1980

Qy  1981  ACTATGATGACTTGTGCTACGAGAACACCCTCACTGTGAAGTTGGCCTCATTGTTGCCACC  2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  ACTATGATGACTTGTGCTACGAGAACACCCTCACTGTGAAGTTGGCCTCATTGTTGCCACC  

Qy  2041  GGAAGCAACGCCTGCTACATGGTAATGCTGAAGAGATGCGTAATGTGGAGCTGTGACGGAGGAG  2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  GGAAGCAACGCCTGCTACATGGTAATGCTGAAGAGATGCGTAATGTGGAGCTGTGACGGAGGAG  2100

Qy  2101  GGACGGATGTGTGTCAACATGGAGTGGGAGCATTGGGACAATGGCTGCCTGGATGAC    2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  GGACGGATGTGTGTCAACATGGAGTGGGAGCATTGGGACAATGGCTGCCTGGATGAC    2160

Qy  2161  TTGCGGACCGTGTTTGATGTTGCTGTGGATGAGCTTTCTCAACCCTGGCAAACAGAGG    2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTGCGGACCGTGTTTGATGTTGCTGTGGATGAGCTTTCTCAACCCTGGCAAACAGAGG    2220

Qy  2221  TTCGAGAAGATGATCAGCGGCATGCGGCATTGTACTTGGAGAGATTGTGCCAACATTCTCATCGAT  2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  TTCGAGAAGATGATCAGCGGCATGCGGCATTGTACTTGGAGAGATTGTGCCAACATTCTCATCGAT  2280

Qy  2281  TTCACGAAGCGGGGCGTCTCTTCCGAGCGCATCTCAGAGCGCCTCAAGACAAGGGGA    2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TTCACGAAGCGGGGCGTCTCTTCCGAGCGCATCTCAGAGCGCCTCAAGACAAGGGGA    2340

Qy  2341  ATCTCTGAAACTAAGTTCCTGTCTCAGATAGAGAGCGACTGCCTAGCCCTGCTACAGGTT  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  ATCTCTGAAACTAAGTTCCTGTCTCAGATAGAGAGCGACTGCCTAGCCCTGCTACAGGTT  2400

Qy  2401  CGTGCCATCCTGCGCGCACCTAGGGCTGGAGAGCACGTGCGATGACAGCAGCATCATCGTGAAG  2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CGTGCCATCCTGCGCGCACCTAGGGCTGGAGAGCACGTGCGATGACAGCAGCATCATCGTGAAG  2460

Qy  2461  GAGGTGTGCACTGTGGTTGCCGGCGCTGCACAGCTCTGTGGCGCGCATGGCCGCC     2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  GAGGTGTGCACTGTGGTTGCCGGCGCTGCACAGCTCTGTGGCGCGCATGGCCGCC     2520

Qy  2521  GTAGTGGACAAGATAAGAGAGAACCGTGGGCTGGGCTGACAACCCCAAAGTGACAGTGGGCGTG  2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  GTAGTGGACAAGATAAGAGAGAACCGTGGGCTGGGCTGACAACCCCAAAGTGACAGTGGGCGTG  2580

Qy  2581  GACGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATGCATGAGACCGTGAGA   2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  GACGGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATGCATGAGACCGTGAGA   2640

Qy  2641  GATCTGGCTCCGAAATGTGACGTGTCCTTCCTGGAGGAATCCTGGAATCCGAAGGGA     2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  GATCTGGCTCCGAAATGTGACGTGTCCTTCCTGGAGGAATCCTGGAATCCGAAGGGA     2700

Qy  2701  GCAGCTCTCATCACTGCCGTGGCCTGCCCATCCGGAGGCTGGGCAGGATAG  2754
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GCAGCTCTCATCACTGCCGTGGCCTGCCCATCCGGAGGCTGGGCAGGATAG  2754

Search completed: May 27, 2003, 19:22:56
Job time : 362 secs
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: May 27, 2003, 16:35:06 ; Search time 361 Seconds
(without alignments)
10073.554 Million cell updates/sec

Title: US-09-808-743A-1
Perfect score: 2754
Sequence: 1 atgatcgcctcgcatatgat.........gggaggctggcagagatag 2754

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 828747 seqs, 660231138 residues

Word size : 50

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Match	Length	DB	ID	Description
1	2754	100.0	2754	10	US-09-808-743-1		Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-808-743-1
; Sequence 1, Application US/09808743
; Patent No. US20020068711A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: PEDERSEN, Peter
; APPLICANT: MATHUPALA, Saroj
; TITLE OF INVENTION: ARREST OF PROLIFERATION OF HIGHLY GLYCOLYTIC TUMORS
; FILE REFERENCE: JHU1720-1
; CURRENT APPLICATION NUMBER: US/09/808,743
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/189,222

; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2754
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-808-743-1

Query Match   100.0%; Score 2754; DB 10; Length 2754;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2754; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG 60

Qy    61 AAGGTTGACCAATTTCTCTACCACATGCGTCTCTCAGATGAGACCCTTCTGCGAGATTTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AAGGTTGACCAATTTCTCTACCACATGCGTCTCTCAGATGAGACCCTTCTGCGAGATTTCT 120

Qy   121 AGGCGGTTCCGGAAGGAGATGGAGAAAGGGCTAGGAGCTACCACGCACCCTACAGCAGCT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGGCGGTTCCGGAAGGAGATGGAGAAAGGGCTAGGAGCTACCACGCACCCTACAGCAGCT 180

Qy   181 GTGAAAATGTTGCCTACCTTTGTGAGGTCAACTCCGGATGGGACAGAACATGGGGAGTTC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGAAAATGTTGCCTACCTTTGTGAGGTCAACTCCGGATGGGACAGAACATGGGGAGTTC 240

Qy   241 CTGGCTCTGGATCTTGGAGGAACCAACTTCCGTGCTCCGAGTAAGGTGACGGACAAT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTGGCTCTGGATCTTGGAGGAACCAACTTCCGTGCTCCGAGTAAGGTGACGGACAAT 300

Qy   301 GGCCTCCAGAGAGTGGAGATGGAAACCAGATCTACGCCATCCTCTTGAGGACATCATCGCGG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGCCTCCAGAGAGTGGAGATGGAAACCAGATCTACGCCATCCTCTTGAGGACATCATCGCGG 360

Qy   361 GGCAGTGGAACCCAGCTGTTTGACCACATGCCTGGCCAACTTCATGGACAAG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCAGTGGAACCCAGCTGTTTGACCACATGCCTGGCCAACTTCATGGACAAG 420

Qy   421 CTACAAATCAAAGAGAAGAGAGTTTTTGGTCTCGTGGACTAAGGGGTTCAAGTCCAGTGGCGTG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTACAAATCAAAGAGAAGAGAGTTTTTGGTCTCGTGGACTAAGGGGTTCAAGTCCAGTGGCGTG 480

Qy   481 ACAAAACTGGATGAGAGAGTTTTTGGTCTCGTGGACTAAGGGGTTCAAGTCCAGTGGCGTG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACAAAACTGGATGAGAGAGTTTTTGGTCTCGTGGACTAAGGGGTTCAAGTCCAGTGGCGTG 540

Qy   541 GAAGGCAGAGATGGTGGCGGTGGAACCTTGATCCGGAAGGTTATCCAGGCGAGGGGACTTTGAC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GAAGGCAGAGATGGTGGCGGTGGAACCTTGATCCGGAAGGTTATCCAGGCGAGGGGACTTTGAC 600

Qy   601 ATTGACATTGTGGCCGTGGTGAATGACACAGTTGGGACCCATGATGACTTGTGGCTATGAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATTGACATTGTGGCCGTGGTGAATGACACAGTTGGGACCCATGATGACTTGTGGCTATGAT 660

Qy   661 GATCAGAACTGCCAGAGATTGGTCTCATTGTGGCACTGGCAGCCAACGCCTGCTACATGGAG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GATCAGAACTGCCAGAGATTGGTCTCATTGTGGCACTGGCAGCCAACGCCTGCTACATGGAG 720

Qy   721 GAAATGCGTCATATTGACATGGTCGGAGGGAGATGAGGGGCCATGTGCATCAACATGGAG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAAATGCGTCATATTGACATGGTCGGAGGGAGATGAGGGGCCATGTGCATCAACATGGAG 780

Qy   781 TGGGGAGCCTTTGGGGACGACGTACACTCAATGACATCCGAACCGAGTTGACCGAGAG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TGGGGAGCCTTTGGGGACGACGTACACTCAATGACATCCGAACCGAGTTGACCGAGAG 840

Qy   841 ATCGACATGGCCTCGCTGAACCCTGGAAGCAGCAGCTGTTTGAGAAGATGATTAGCGGGATG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 ATCGACATGGCCTCGCTGAACCCTGGAAGCAGCAGCTGTTTGAGAAGATGATTAGCGGGATG 900
```

Wed May 28 08:37:51 2003

us-09-808-743a-1.oligo.rni

```
Db  1818 CTGCTGGTGCGTGTCGTAATGCAAGCGGAGGGCGTGGAGATGCATAACAAGATCTAC 1877

Qy  1681 TCCATCCCACAGAGGAGGTTATGCATGGCCACTGGGGAAGAGCTCTTCGACCACATTGTCCAG 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1878 TCCATCCCACAGAGGAGGTTATGCATGGCCACTGGGGAAGAGCTCTTCGACCACATTGTCCAG 1937

Qy  1741 TGCATTGCGGACTTCCTGGAGTACATGGCCATGAAGGGCGTGTCCCTGCCTTTGGGTTTC 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1938 TGCATTGCGGACTTCCTGGAGTACATGGCCATGAAGGGCGTGTCCCTGCCTTTGGGTTTC 1997

Qy  1801 ACATTCTCCTTCCCTTGCCAGCAGAACAGCCTAGACCAGGCATCCTCCTCAAGTCGGACA 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1998 ACATTCTCCTTCCCTTGCCAGCAGAACAGCCTAGACCAGGCATCCTCCTCAAGTCGGACA 2057

Qy  1861 AAGGGATTCAAGGCATCTGGCTGCGAGGGTGAGGATGGATGTGGTCACCTTGCTGAAGGAAGCG 1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2058 AAGGGATTCAAGGCATCTGGCTGCGAGGGTGAGGATGGATGTGGTCACCTTGCTGAAGGAAGCG 2117

Qy  1921 ATTCACCCGGCGAGAGGAGTTTGACCTGGATGGTGGTTGCCGTGCTGAATGACACAGTTGGG 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2118 ATTCACCCGGCGAGAGGAGTTTGACCTGGATGGTGGTTGCCGTGCTGAATGACACAGTTGGG 2177

Qy  1981 ACTATGATGACTTGTGGCTACGAAGACCCTCACTGTGAAGTTGGCCTCATTGTTGCCACC 2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2178 ACTATGATGACTTGTGGCTACGAAGACCCTCACTGTGAAGTTGGCCTCATTGTTGCCACC 2237

Qy  2041 GGAAGCAACGCCTGCTACATGGAAGAGATGCGTAATGTGGAGCTGTGGACGGAGGAGGAG 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2238 GGAAGCAACGCCTGCTACATGGAAGAGATGCGTAATGTGGAGCTGTGGACGGAGGAGGAG 2297

Qy  2101 GGACGGATGTGTGTCAACATGGAGTGGGGAGCATTTGGGGACAATGGCTGCCTGCGATGAC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2298 GGACGGATGTGTGTCAACATGGAGTGGGGAGCATTTGGGGACAATGGCTGCCTGCGATGAC 2357

Qy  2161 TTGCGGACCGTGTTTGATGTTGCTGTGGATGAGCTTCTCTCAACCCTGCAAACAGAGG 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2358 TTGCGGACCGTGTTTGATGTTGCTGTGGATGAGCTTCTCTCAACCCTGCAAACAGAGG 2417

Qy  2221 TTCGAGAAGATGATCAGCGGCATGTACTTGGGAGAGATTGTGCGCAACATTCTCATCGAT 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2418 TTCGAGAAGATGATCAGCGGCATGTACTTGGGAGAGATTGTGCGCAACATTCTCATCGAT 2477

Qy  2281 TTCACGGAAGCGGGGGCGGCTGCTCTTCCGAGCCGCATCTCAGAGCGCCTCAAGACAAGGGGA 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2478 TTCACGGAAGCGGGGGCGGCTGCTCTTCCGAGCCGCATCTCAGAGCGCCTCAAGACAAGGGGA 2537

Qy  2341 ATCTCTGAAACTAAGTTCCTGTCTCAGATAGAGAGCGACTGCCTAGCCCTGCTACAGGTT 2400
         ||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2538 ATCTTTGAAACTAAGTTCCTGTCTCAGATAGAGAGCGACTGCCTAGCCCTGCTACAGGTT 2597

Qy  2401 CGTGCCATCCTGCCGCACCTAGGGCTGGAGCTGGAGAGCACGTGCGATGACAGCATCATCGTGAAG 2460
         |||||||||||||||||||||||||||||||||||||||||||  |||||||||||||||||||
Db  2598 CGTGCCATCCTGCCGCACCTAGGGCTGGAGCTGGAGAGCACGTGCGATGACAGCATCATCGTGAAG 2657

Qy  2461 GAGGTGTGCACTGTGGTTCGCCCGGCCGCTCGTGTTGGCGCAGCTCTGTGGCCGCCCC 2520
         ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  2658 GAGGTGTGCACTGTGGTTCGCCCGGCCGCTCGTGTTGGCGCAGCTCTGTGGCCGCCGC 2717

Qy  2521 GTAGTGGACAAGATAAGAGAGAACCGTGGGCTGGACAACCCCAAAGTGACAGTGGGCGTG 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2718 GTAGTGGACAAGATAAGAGAGAACCGTGGGCTGGACAACCTCAAAGTGACAGTGGGCGTG 2777

Qy  2581 GACCGGGACTTCTGTATAAAGCTTCATCCTCACTTTGCCAAGGTCATGCATGAGACGGTGAGA 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2778 GACCGGGACTTCTGTATAAAGCTTCATCCTCACTTTGCCAAGGTCATGCATGAGACGGTGAGA 2837

Qy  2641 GATCTGGCTCCGAAATGTCACGTGTCCTTCCTTGGAATCCGAGGACGCAGTGGGAAGGGA 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2838 GATCTGGCTCCGAAATGTCACGTGTCCTTCCTTGGAATCCGAGGACGCAGTGGGAAGGGA 2897

Qy  2701 GCAGCTCTCATCACTGCCGCTGCCCGCATCCGGGAGGCTGGGCAGCAGAGATAG 2754
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2898 GCAGCTCTCATCACTGCCGCTGCCCGCATCCGGGAGGCTGGGCAGCAGAGATAG 2951
```

CORRESPONDENCE ADDRESS:
  ADDRESSEE: Arnold, White & Durkee
  STREET: P.O. Box 4433
  CITY: Houston
  STATE: TX
  COUNTRY: US
  ZIP: 77210
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/588,976
  FILING DATE: Concurrently herewith
  CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
  NAME: Fussey, Shelley P.M.
  REGISTRATION NUMBER: 39,458
  REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (512) 418-3000
  TELEFAX: (512) 474-7577
  TELEX: n/a
INFORMATION FOR SEQ ID NO: 15:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 3635 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear

US-08-588-976-15

Query Match   92.6%; Score 2550; DB 2; Length 3635;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2750; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   1 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAGTGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAGTGCAG 257

Qy   61 AAGGTTGACCAATTTCTCTACCACATGGCGTCTCTCAGATGAGACCCTTCTGGAGATTTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258 AAGGTTGACCAATTTCTCTACCACATGGCGTCTCTCAGATGAGACCCTTCTGGAGATTTCT 317

Qy   121 AGGCCGGTTCCGGAAGGAGATGAGAAAGGGCTAGGAGGCTACCACGCACCCTACAGCAGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318 AGGCCGGTTCCGGAAGGAGATGAGAAAGGGCTAGGAGGCTACCACGCACCCTACAGCAGCT 377

Qy   181 GTGAAAATGTTGCCTACCTTTGTGAGGTCAACTCCGGATGGACAGAACATGGGGAGGTTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   378 GTGAAAATGTTGCCTACCTTTGTGAGGTCAACTCCGGATGGACAGAACATGGGGAGGTTC 437

Qy   241 CTGGCTCTGATCTTGGAGAACCAACTTCCGTGTCCTCCGTGTCCTCCGAGTAAGGTGACGGACAAT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   438 CTGGCTCTGATCTTGGAGAACCAACTTCCGTGTCCTCCGTGTCCTCCGAGTAAGGTGACGGACAAT 497

Qy   301 GGCCTCCAGAGAGTGGAGATGGAGAACCAGATCTACGCCATCCTTGAGGACATCATGCGG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   498 GGCCTCCAGAGAGTGGAGATGGAGAACCAGATCTACGCCATCCTTGAGGACATCATGCGG 557

Qy   361 GGCAGTGGAACCCAGCTGTTGACCACATCCCGGCCAACTTCATGGACAAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   558 GGCAGTGGAACCCAGCTGTTGACCACATCCCGGCCAACTTCATGGACAAG 617

Qy   421 CTACAAATCAAAGAAGAGAAGAAGCTCCCCTGGTTTCACCTTCTCGTTCCCCACCAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   618 CTACAAATCAAAGAAGAGAAGAAGCTCCCCTGGTTTCACCTTCTCGTTCCCCACCAG 677

Qy   481 ACAAAACTGGATGGAGAGTTTTTGGTCTCGTTGGGACTAAGGGGTTCAAGTCCAGTGGCGTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   678 ACAAAACTGGATGGAGAGTTTTTGGTCTCGTTGGGACTAAGGGGTTCAAGTCCAGTGGCGTG 737

Qy   541 GAAGGCAGAGATGTGGTGGACCTGATCCGGAAGGTTATCCAGCAGAGGGGACTTTGAC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   738 GAAGGCAGAGATGTGGTGGACCTGATCCGGAAGGTTATCCAGCAGAGGGGACTTTGAC 797

Qy   601 ATTGACATTGTGGCCGTGGCCGTGGCCATGGCCACTGGCACAACGCCTGCTACATGGAG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   798 ATTGACATTGTGGCCGTGGCCGTGGCCATGGCCACTGGCACAACGCCTGCTACATGGAG 857

Qy   661 GATCAGAACTGCGAGATTGGTCTCATTGTGGCACTGGCAGCACAACGCCTGCTACACATGGAG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   858 GATCAGAACTGCGAGATTGGTCTCATTGTGGCACTGGCAGCACAACGCCTGCTACACATGGAG 917

Qy   721 GAAATGCGTCATATTGACATGTGGAGGGAGTGAGGGGCCATGTGCATCAACATGGAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   918 GAAATGCGTCATATTGACATGTGGAGGGAGTGAGGGGCCATGTGCATCAACATGGAG 977

Qy   781 TGGGAGCCTTTGGGGACGACGGTACACTCAATGACATCCGAACCGAGTTTGACCGAGAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   978 TGGGAGCCTTTGGGGACGACGGTACACTCAATGACATCCGAACCGAGTTTGACCGAGAG 1037

Qy   841 ATCGACATGGGCTCGCTGAACCCTGGGAAGCAGCTGTTTGAGAAGTGATTAGCGGGATG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038 ATCGACATGGGCTCGCTGAACCCTGGGAAGCAGCTGTTTGAGAAGTGATTAGCGGGATG 1097

Qy   901 TACATGGGGAGCTGGTCAGGCTCATCCTGGTGAAGATGGCCAAGGCAGAGAGCTGTTCTTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1098 TACATGGGGAGCTGGTCAGGCTCATCCTGGTGAAGATGGCCAAGGCAGAGAGCTGTTCTTC 1157

Qy   961 CAAGGGAAACTCAGCCCAGAACTCCTTACCACTGGCTCCTTCGAGACCAAAGATGTCTCG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1158 CAAGGGAAACTCAGCCCAGAACTCCTTACCACTGGCTCCTTCGAGACCAAAGATGTCTCG 1217

Qy   1021 GATATTGAAGAGGATAAGGATGGAATCGAGAAGGCCTACCAAAATCCTGATGCGCCTGGGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1218 GATATTGAAGAGGATAAGGATGGAATCGAGAAGGCCTACCAAAATCCTGATGCGCCTGGGT 1277

Qy   1081 CTGAATCCATTGCAGGAGGATTGTGTGGCCACGCCACGGAATCTGCCAGATTGTCCACG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1278 CTGAATCCATTGCAGGAGGATTGTGTGGCCACGCCACGGAATCTGCCAGATTGTCCACG 1337

Qy   1141 CGCTCGGCCAGTCTGTGCGCAGCCACCTGGCCGCGGTGCTGTGGCGAATCAAAGAGAAC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1338 CGCTCGGCCAGTCTGTGCGCAGCCACCTGGCCGCGGTGCTGTGGCGAATCAAAGAGAAC 1397

Qy   1201 AAGGGCGAGGAGGACTTCGGTCCACCATCGGTTCGTCCGTCTCACAGAGAAACAT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1398 AAGGGCGAGGAGGACTTCGGTCCACCATCGGTTCGTCCGTCTCACAGAGAAACAT 1457

Qy   1261 CCCCATTTGCCAAGCGTCTCCATAAGGCAGTGAGGAGCTGGTGCCCGACTGTGATGTC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1458 CCCCATTTGCCAAGCGTCTCCATAAGGCAGTGAGGAGCTGGTGCCCGACTGTGATGTC 1517

Qy   1321 CGCTTCCTCCGCTCTGAGGATGGCAGCGGGGCTGCGCTATGGTGACGGCGGTGGCT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1518 CGCTTCCTCCGCTCTGAGGATGGCAGCGGGGCTGCGCTATGGTGACGGCGGTGGCT 1577

Qy   1381 TACCGTCTGCTGACCAACACCGGCCCCGCAGAGACCCGGAGTCTCTGAAGCTGAGC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1578 TACCGTCTGCTGACCAACACCGGCCCCGCAGAGACCCGGAGTCTCTGAAGCTGAGC 1637

Qy   1441 CACGAGCAGCTTCTGGAGGTTAAGAAGAATGAAGGTGAAATGGAGCAGGGTCTGAGC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1638 CACGAGCAGCTTCTGGAGGTTAAGAAGAATGAAGGTGAAATGGAGCAGGGTCTCGAGC 1697

Qy   1501 AAGGAGACGCACAGAGAAGGAGACTTCTTGGCCTTGGATCTTGGAGAACACTTCCCGGGTC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1698 AAGGAGACGCACAGAGAAGGAGACTTCTTGGCCTTGGATCTTGGAGAACAACTTCCCGGGTC 1757

Qy   1561 GATGGCACAGAGAAAGGAGACTTCTTGGCCTTGGATCTTGGAGAGCAAACTTCCCGGGTC 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1758 GATGGCACAGAGAAAGGAGACTTCTTGGCCTTGGATCTTGGAGAGCAAACTTCCCGGGTC 1817

Qy   1621 CTGCTGCTGTGCGTGTGCGTAATGCGAGGGGCTGCTAATGCATAACAAGATCTAC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1680

Qy   721 GAAATGCTCATATTGACATGGTCGAGGAGATGAGGGGCATGCATCAACATGGAG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   918 GAAATGCTCATATTGACATGGTCGAGGAGATGAGGGGCATGCATCAACATGGAG 977

Qy   781 TGGGAGCCTTTGGGGACGACGTACACTCAATGACATCCGAACCGAGTTTGACCGAGAG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   978 TGGGAGCCTTTGGGGACGACGTACACTCAATGACATCCGAACCGAGTTTGACCGAGAG 1037

Qy   841 ATCGACATGGGCTGCTCGGAAGCAGCTGTTTGAGAGATGATTAGCGGGATG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038 ATCGACATGGGCTGCTCGGAAGCAGCTGTTTGAGAGATGATTAGCGGGATG 1097

Qy   901 TACATGGGGGAGCTGGTCAGGCTCATCCTGGTGAAGATGGCCAAGGCAGAGCTGTTGTTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1098 TACATGGGGGAGCTGGTCAGGCTCATCCTGGTGAAGATGGCCAAGGCAGAGCTGTTGTTC 1157

Qy   961 CAAGGGAAACTCAGCCCAGAAACTTCCTTACCACTGGCCTCCTTCGAGACCCAAAGATGTCTCG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1158 CAAGGGAAACTCAGCCCAGAAACTTCCTTACCACTGGCCTCCTTCGAGACCCAAAGATGTCTCG 1217

Qy   1021 GATATTGAAGAGGATAAGGATGAATCGAGAAGCCTACCAAATCTGATGCGCCTGGGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1218 GATATTGAAGAGGATAAGGATGAATCGAGAAGCCTACCAAATCTGATGCGCCTGGGT 1277

Qy   1081 CTGAATCCATTGCAGCTGTGTGCGCCACGCACCGAATCTGCCAGATTGTGTCCACG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1278 CTGAATCCATTGCAGCTGTGTGCGCCACGCACCGAATCTGCCAGATTGTGTCCACG 1337

Qy   1141 CGCTCGGCCCAGTCTGTGCGCAGCCACCCTGGCCGCGTGGTGGTCCGCGACTTCGGTCTCGTGTGGGT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1338 CGCTCGGCCCAGTCTGTGCGCAGCCACCCTGGCCGCGTGGTGGTCCGCGACTTCGGTCTCGTGTGGGT 1397

Qy   1201 AAGGGCGGAGGAGCGACTTCGCTCGCACCATCGGTGTCGATGGCTGCATGGCGCAAGGGCCAAGGGGCT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1398 AAGGGCGGAGGAGCGACTTCGCTCGCACCATCGGTGTCGATGGCTGCATGGCGCAAGGGCCAAGGGGCT 1457

Qy   1261 CCCCATTTGTCCAAGCGTCTCTGAGGATGGCAGGCGGCAAGGGCGGTCGCGTGGGCGGTGGCT 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1458 CCCCATTTGTCCAAGCGTCTCTGAGGATGGCAGGCGGCAAGGGCGGTCGCGTGGGCGGTGGCT 1517

Qy   1321 CGCTTCCTCCCGCCTCTGAGGATGGCACCGGCCCGTCGCAGAGACCCTGAAAGCTGAGC 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1518 CGCTTCCTCCCGCCTCTGAGGATGGCACCGGCCCGTCGCAGAGACCCTGAAAGCTGAGC 1577

Qy   1381 TACCGTCTGGCTGACCAACACCGGGAGCGCGGCCGCGAGTGTGTGCCACTCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1578 TACCGTCTGGCTGACCAACACCGGGAGCGCGGCCGCGAGTGTGTGCCACTCCA 1637

Qy   1441 CACGAGCACGAGAAAGCGATGCGGCCCGTGTGAAGATGAAGGTGAAATGGAGCAGGGTCTGAGC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1638 CACGAGCACGAGAAAGCGATGCGGCCCGTGTGAAGATGAAGGTGAAATGGAGCAGGGTCTGAGC 1697

Qy   1501 AAGGAGACGACGTGCGGTCGCCCCTGTGAAGATGTGCCCACTTACGTGTGTGCCACTCCA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1698 AAGGAGACGACGTGCGGTCGCCCCTGTGAAGATGTGCCCACTTACGTGTGTGCCACTCCA 1757

Qy   1561 GATGGCACAGAGAAAGGAGACTTCTTGGAGATCTTGGAGGAACAAACTTCCGGGTC 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1758 GATGGCACAGAGAAAGGAGACTTCTTGGAGATCTTGGAGGAACAAACTTCCGGGTC 1817

Qy   1621 CTGCTGGTGCGTGCGTAATGGCAAGGCGAGGGCGTGAAGCTCTTTCGCACCCACATTGTCCAG 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1818 CTGCTGGTGCGTGCGTAATGGCAAGGCGAGGGCGTGAAGCTCTTTCGCACCCACATTGTCCAG 1877

Qy   1681 TCCATCCCACAGGAGGTTATGCACTGGGCAACGCTCTTCGAACACGCTCTTCGACCACATTGTCCAG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1878 TCCATCCCACAGGAGGTTATGCATGGGCAACGCTCTTCGAACACGCTCTTCGACCACATTGTCCAG 1937

Qy   1741 TGCATTGCGGACTTCCTGCAGTACATGGGCATGAAGAGCCGTGTCCCTTGCCTTTGGGTTTC 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1938 TGCATTGCGGACTTCCTGCAGTACATGGGCATGAAGAGCCGTGTCCCTTGCCTTTGGGTTTC 1997

Qy   1801 ACATTCTCCTTCCCTTGCCCAGCAGAACAGCCTAGACCAGAACATCCTCCTCAAGTGGACA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1998 ACATTCTCCTTCCCTTGCCCAGCAGAACAGCCTAGACCAGAACATCCTCCTCAAGTGGACA 2057

Qy   1861 AAGGGGATTCAAGGCATCTGGCTGCCGAGGTGAGGGTGAAGGTGGTCACCTTGCTGACGGAGGAGCG 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2058 AAGGGGATTCAAGGCATCTGGCTGCCGAGGTGAGGGTGAAGGTGGTCACCTTGCTGACGGAGGAGCG 2117

Qy   1921 ATTCACCGGCGAGAGAGTTTGACCTGCTGGTCGTCCTGTGTGAATGACACAGTTGGG 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2118 ATTCACCGGCGAGAGAGTTTGACCTGCTGGTCGTCCTGTGTGAATGACACAGTTGGG 2177

Qy   1981 ACTATGATGACTTGTGGCTACGAAGACCCTCACTGTGAAGTTGGCCTCATTGTTGGCACC 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2178 ACTATGATGACTTGTGGCTACGAAGACCCTCACTGTGAAGTTGGCCTCATTGTTGGCACC 2237

Qy   2041 GGAAGCAACGCCTGCTACACTGGAAGAGATGCGTAATGCTAATGTGGTGGTGGACGGAGGAG 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2238 GGAAGCAACGCCTGCTACACTGGAAGAGATGCGTAATGCTAATGTGGTGGTGGACGGAGGAG 2297

Qy   2101 GGACGGATGTGTGTCAACATGGACTGGGAGCATTGGGGAGCAATGCTGCCTGGATGAC 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2298 GGACGGATGTGTGTCAACATGGACTGGGAGCATTGGGGAGCAATGCTGCCTGGATGAC 2357

Qy   2161 TTGCGGACCGTGTTGATGTGTGTGCTGTGGATGAGCTTTCTCTCAACCCTGGCAAACAGAGG 2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2358 TTGCGGACCGTGTTGATGTGTGTGCTGTGGATGAGCTTTCTCTCAACCCTGGCAAACAGAGG 2417

Qy   2221 TTCGAGAAGATGATCAGCGGCATGTACTTGGGAGAGATTGTGCGCAACATTCTCATCGAT 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2418 TTCGAGAAGATGATCAGCGGCATGTACTTGGGAGAGATTGTGCGCAACATTCTCATCGAT 2477

Qy   2281 TTCACCAAGCGCGGGGCTGCTCTTCCGAGGCCGCATCTCAGAGGCCGCCTCAAGACAACAGGGGA 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2478 TTCACCAAGCGCGGGGCTGCTCTTCCGAGGCCGCATCTCAGAGGCCGCCTCAAGACAACAGGGGA 2537

Qy   2341 ATCTCTGAAACTAAGTTCCTGCGCCACCTAGGGCTGGAGAGAGCGACTGCTGCTACCAGGTT 2400
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2538 ATCTTTGAAACTAAGTTCCTGTCTCGTCAGATAGAGAGCGACTGCTGCTACCAGGTT 2597

Qy   2401 CGTGCCATCTCGCCGCCACCTAGGGCTGGAGAGCACGTGCAGTGCGCAGCATCATCGTGAAG 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2598 CGTGCCATCTCGCCGCCACCTAGGGCTGGAGAGCACGTGCAGTGCGCAGCATCATCGTGAAG 2657

Qy   2461 GAGGTGTGCACTGTGGTTGCCCGGCGCGTCTCTGAAGCTGAGC 2520
          |||||||||||||||||||||||||||||||||||||||||||
Db   2658 GAGGTGTGCACTGTGGTTGCCCGGCGCGTCTGTGGCGCATGGCCGCC 2717

Qy   2521 GTAGTGGACAAGATAAGAGGAGAACCGTGGGCTGACACACCCAAAGTGACAGTGGGCGTG 2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2718 GTAGTGGACAAGATAAGAGGAGAACCGTGGGCTGACACACCCAAAGTGACAGTGGGCGTG 2777

Qy   2581 GACCGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATGCATGGCCGCCGC 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2778 GACCGGACTCTGTATAAGCTTCATCCTCACTTTGCCAAGGTCATGCATGGCCGCCGC 2837

Qy   2641 GATCTGGCTCCGAAATGTCGACGTTCATCCTCACTTTGCCAAGGTCATGCAGGACGTGGGAAGGGA 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2838 GATCTGGCTCCGAAATGTCGACGTTCATCCTCACTTTGCCAAGGTCATGCAGGACGTGGGAAGGGA 2897

Qy   2701 GCAGCTCTCATCACTGCCGCCTGGCCGCCGCATCCGGAGGCTGGCAGAGATAG 2754
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2898 GCAGCTCTCATCACTGCCGCCTGGCCGCCGCATCCGGAGGCTGGCAGAGATAG 2951

RESULT 2
US-08-588-976-15
; Sequence 15, Application US/08588976
; Patent No. 5891717
; GENERAL INFORMATION:
;   APPLICANT: Christopher B. Newgard, et al.
;   TITLE OF INVENTION: Methods and Compositions for
;   TITLE OF INVENTION: Inhibiting Hexokinase
;   NUMBER OF SEQUENCES: 43

OM nucleic - nucleic search, using sw model

```
                    GenCore version 5.1.4_p5_4578
                    Copyright (c) 1993 - 2003 Compugen Ltd.
```

Run on: May 27, 2003, 13:52:56 ; Search time 101 Seconds
                                        (without alignments)
                                        8362.255 Million cell updates/sec

Title: US-09-808-743A-1
Perfect score: 2754
Sequence: 1 atgatcgcctcgcatatgat........gggaggctgggcagagatag 2754

Scoring table: OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size :   50

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   Issued_Patents_NA:*
                1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

            Pred. No. is the number of results predicted by chance to have a
            score greater than or equal to the score of the result being printed,
            and is derived by analysis of the total score distribution.

                              SUMMARIES

                 Query
Result   Score   Match Length DB ID   Description
No.
------- -----   ------------ -- --   -----------
   1   2550   92.6 3635   2   US-08-588-983-15   Sequence 15, Appl
   2   2550   92.6 3635   2   US-08-588-976-15   Sequence 15, Appl

                             ALIGNMENTS

RESULT 1
US-08-588-983-15
; Sequence 15, Application US/08588983
; Patent No. 5854067
; GENERAL INFORMATION:
; APPLICANT: Christopher B. Newgard, et al.
; TITLE OF INVENTION: Methods and Compositions
; TITLE OF INVENTION: for Inhibiting Hexokinase
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Arnold, White & Durkee
;   STREET: P.O. Box 4433
;   CITY: Houston
;   STATE: TX
;   COUNTRY: US
;   ZIP: 77210
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
Query Match                92.6%;   Score 2550;   DB 2;   Length 3635;
Best Local Similarity      99.9%;   Pred. No. 0;
Matches 2750;   Conservative   0;   Mismatches  4;   Indels  0;   Gaps  0;

Qy    1 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  198 ATGATCGCCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG  257

Qy   61 AAGGTTGACCAATTTCTCTACCACATGACCATGACCATGTTCTGAGATTTCT          120
        |||||||||||||||||||||||||||||||||||||||||||||| ||| |||||||
Db  258 AAGGTTGACCAATTTCTCTACCACATGACCATGACCATGTTCTGAGATTTCT          317

Qy  121 AGGGGGTTCCGAAGGAGATGGAGAAAGGCTAGGAGCTACCACGCACCCTACAGCAGCT     180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  318 AGGGGGTTCCGAAGGAGATGGAGAAAGGCTAGGAGCTACCACGCACCCTACAGCAGCT     377

Qy  181 GTGAAAATGTTGCCTACCTTTGTGAGGTCAACTCCGATGGGACAGAACATGGGGGAGTTC  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  378 GTGAAAATGTTGCCTACCTTTGTGAGGTCAACTCCGATGGGACAGAACATGGGGGAGTTC  437

Qy  241 CTGGCTCTGGATCTTGGAGGAACCAACTTCCGTGCTCCGAGTAAGGGTGACGGACAAT    300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  438 CTGGCTCTGGATCTTGGAGGAACCAACTTCCGTGTGCTCCGAGTAAGGGTGACGGACAAT  497

Qy  301 GGCCTCCAGAGAGTGGAGATGGAGAACCAGATCTACGCCATCCTTGAGGACATCATGCGG  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  498 GGCCTCCAGAGAGTGGAGATGGAGAACCAGATCTACGCCATCCTGAGGACATCATGCGG  557

Qy  361 GGCAGTGGAACCCAGCTGTTTGACCACATCGCCGAATGCCTGGCCAACTTCATGGACAAG  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  558 GGCAGTGGAACCCAGCTGTTTGACCACATCGCCGAATGCCTGGCCAACTTCATGGACAAG  617

Qy  421 CTACAAATCAAAGAGAAGAAGAAGCTCCCTCTGGGTTTCACCTTCGTTCCCCTGCCACCAG  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  618 CTACAAATCAAAGAGAAGAAGAAGCTCCCTCTGGGTTTCACCTTCGTTCCCCTGCCACCAG  677

Qy  481 ACAAAACTGGATGAGAGTTTTTGGTCTCGTGGACTAAGGGGTTCAAGTCCAGTGCCGTG  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  678 ACAAAACTGGATGAGAGTTTTTTGGTCTCGTGGACTTAAGGGGTTCAAGTCCAGTGCCGTG  737

Qy  541 GAAGGCAGAGATGTGGTGGACCTGATCCGGAAGGTTATCCAGCGCAGAGGGGACTTTGAC  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  738 GAAGGCAGAGATGTGGTGGACCTGATCCGGAAGGCTATCCAGCGCAGAGGGGACTTTGAC  797

Qy  601 ATTGACATTGTGCCCGTGGTGAATGACACAGTTGGGACCACAGTTGGTGGCTATGAT     660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  798 ATTGACATTGTGCCCGTGGTGAATGACACAGTTGGGACCACAGTTGGTGGCTATGAT     857

Qy  661 GATCAGAACTGCCAGATTGGTCTCATTGTGGGCACTGGGCAGCAACGCCGCTACATGGGAG  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  858 GATCAGAACTGCCAGATTGGTCTCATTGTGGGCACTGGGCAGCAACGCCGCTACATGGGAG  917
```

```
FT
FT   protein_bind   /note= "Transcription factor binding site"
FT                  3162..3167
FT                  /*tag= an
FT                  /label= Myb
FT   protein_bind   /note= "Transcription factor binding site"
FT                  3223..3228
FT                  /*tag= ao
FT                  /label= SRE
FT   protein_bind   /note= "Transcription factor binding site"
FT                  3256..3267
FT                  /*tag= ap
FT                  /label= Sp-1,Ap-2,SRE
FT   protein_bind   /note= "Transcription factor binding site"
FT                  3306..3313
FT                  /*tag= aq
FT                  /label= c/ebp
FT   protein_bind   /note= "Transcription factor binding site"
FT                  3352..3360
FT                  /*tag= ar
FT                  /label= NF-IL6
FT   repeat_unit    /note= "Transcription factor binding site"
FT                  3466..3483
FT                  /*tag= as
FT                  /rpt_type= DIRECT
FT                  /label= II
FT   repeat_unit    3492..3509
FT                  /*tag= at
FT                  /rpt_type= DIRECT
FT   protein_bind   /label= III
FT                  3507..3513
FT                  /*tag= au
```

Query Match   2.4%; Score 65; DB 18; Length 5150;
Best Local Similarity 100.0%; Pred. No. 2.7e-21;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 ATGATCGCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4831 ATGATCGCTCGCATATGATCGCCTGCTTATTCACGGAGCTCAACCAAAACCAAGTGCAG 4890

Qy    61 AAGGT   65
         |||||
Db  4891 AAGGT 4895
```

Search completed: May 27, 2003, 14:07:54
Job time : 624 secs

Wed May 28 08:37:51 2003

```
FT
FT
FT   protein_bind   208..220
FT                  /*tag= c
FT                  /label= p53
FT                  /note= "Transcription factor binding site"
FT   protein_bind   231..240
FT                  /*tag= d
FT                  /label= HNF-5,AP-3,c/ebp
FT                  /note= "Transcription factor binding site"
FT   protein_bind   322..327
FT                  /*tag= e
FT                  /label= IRE
FT                  /note= "Transcription factor binding site"
FT   protein_bind   453..459
FT                  /*tag= f
FT                  /label= Myb
FT                  /note= "Transcription factor binding site"
FT   protein_bind   504..523
FT                  /*tag= g
FT                  /label= HNF-5
FT                  /note= "Transcription factor binding site"
FT   protein_bind   529..559
FT                  /*tag= h
FT                  /label= Pea-3,SRE,Ap-2
FT                  /note= "Transcription factor binding site"
FT   repeat_region  605..610
FT                  /*tag= i
FT                  /rpt_type= DIRECT
FT                  /note= "Motif with the potential to form Z-DNA
FT                          structure"
FT   protein_bind   640..647
FT                  /*tag= j
FT                  /label= G1RE,Myc
FT                  /note= "Transcription factor binding site"
FT   protein_bind   723..728
FT                  /*tag= k
FT                  /label= c/ebp
FT                  /note= "transcription factor binding site"
FT   protein_bind   741..746
FT                  /*tag= l
FT                  /label= Pea-3
FT                  /note= "Transcription factor binding site"
FT   protein_bind   899..906
FT                  /*tag= m
FT                  /label= Pea-3
FT                  /note= "Transcription factor binding site"
FT   protein_bind   1037..1043
FT                  /*tag= n
FT                  /label= Ap-1
FT                  /note= "Transcription factor binding site"
FT   protein_bind   1076..1084
FT                  /*tag= o
FT                  /label= HNF-5
FT                  /note= "Transcription factor binding site"
FT   protein_bind   1112..1117
FT                  /*tag= p
FT                  /label= Sp-1
FT                  /note= "Transcription factor binding site"
FT   protein_bind   1414..1419
FT                  /*tag= q
FT                  /label= Pea-3
FT                  /note= "Transcription factor binding site"
FT   protein_bind   1634..1640
FT                  /*tag= r
FT                  /label= PPAR
FT                  /note= "Transcription factor binding site"
FT   protein_bind   1816..1824
FT                  /*tag= s
FT                  /label= Ap-1
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2045..2051
FT                  /*tag= t
FT                  /label= c/ebp,NF-IL6
FT                  /note= "transcription factor binding site"
```

```
FT                  /*tag= u
FT                  /label= Ap-1
FT                  /note= "Transcription factor binding site"
FT   repeat_unit    2090..2099
FT                  /rpt_type= DIRECT
FT                  /label= I
FT   repeat_unit    2133..2142
FT                  /*tag= v
FT                  /rpt_type= DIRECT
FT                  /label= I
FT   protein_bind   2144..2152
FT                  /*tag= w
FT                  /label= Sp-1
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2165..2171
FT                  /*tag= x
FT                  /label= HNF-5
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2325..2332
FT                  /*tag= y
FT                  /label= Ap-2
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2402..2409
FT                  /*tag= z
FT                  /label= Ap-2
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2413..2420
FT                  /*tag= aa
FT                  /label= Ap-1
FT                  /note= "Transcription factor binding site"
FT   repeat_region  2593..2636
FT                  /*tag= ab
FT                  /rpt_type= DIRECT
FT   protein_bind   2757..2766
FT                  /*tag= ac
FT                  /label= p53
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2776..2784
FT                  /*tag= ad
FT                  /label= NF-IL6,Ap-1
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2866..2873
FT                  /*tag= ae
FT                  /label= Ap-2
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2908..2919
FT                  /*tag= af
FT                  /label= Myb,Mbf-1,Mep-1,Mtf-1
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2926..2933
FT                  /*tag= ag
FT                  /label= c/ebp
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2951..2956
FT                  /*tag= ah
FT                  /label= Pea-3
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2988..2993
FT                  /*tag= ai
FT                  /label= Myb
FT                  /note= "Transcription factor binding site"
FT   protein_bind   2998..3003
FT                  /*tag= aj
FT                  /label= Pea-3
FT                  /note= "Transcription factor binding site"
FT   protein_bind   3076..3086
FT                  /*tag= ak
FT                  /label= SRE
FT                  /note= "Transcription factor binding site"
FT   protein_bind   3102..3109
FT                  /*tag= al
FT                  /label= Pea-3
FT                  /note= "Transcription factor binding site"
FT                  /*tag= am
FT                  /label= Ap-2
```